Dear Editor,

We thank the three reviewers for their thoughtful reviews and respond to each point individually below, as well as making changes in the revised manuscript. First, we make three general points, and then respond to each reviewer comment (reviewer comments in bold, our responses in plain text).

1. We asked for additional time to revise the paper to address a key suggestion of reviewer 1, that is, to use a box model to help explain the observed seasonal cycle. Ultimately, we have not included the box model analysis in the paper for the following reasons. As several previous authors (Randerson et al., 2002; Levin et al., 2010) have pointed out, it is difficult to reproduce the ^{14}C bomb spike, seasonal cycle and rate of decline with a simple 4 box model (Northern and Southern Hemispheres, each divided into troposphere and stratosphere). They were right, and we were unable to match the bomb spike peak, timing or interhemispheric offset unless we adjust the transport and flux terms to such an extent that we do not believe it is justifiable to interpret the results in a meaningful way. (For example, to match the maximum bomb peak amplitude in the Northern troposphere, we needed either an unrealistically fast stratosphere-troposphere exchange rate of less than one year or to place 20% of the bomb ^{14}C in the Northern troposphere (rather than stratosphere)). We considered building a more elaborate box model, but concluded that our existing capabilities make it more realistic for us to focus on including ^{14}C a higher resolution global atmospheric transport model for a future publication. Thus, we have substantially revised the discussion around the seasonal cycle to address the reviewer comments and remove sections that are speculative, and we have not included box modelling.
2. Reviewer 2 suggests that there is not sufficient new information or interpretation in the paper to warrant publication in ACP. We respectfully disagree. First, both reviewers 1 and 3 recommend publication with revisions. Second, the Wellington $^{14}CO_2$ record is the longest direct atmospheric record of any trace gas or isotope anywhere in the world, and is the only long-term Southern Hemisphere $^{14}CO_2$ record. It has been used widely and will no doubt continue to be used widely (previous reports on the Wellington record have been directly cited 138 times (Currie et al 2011, Manning et al 1990) and the dataset is the main Southern Hemisphere record used in compiled ^{14}C global records that have been cited more than 500 times (e.g. Hua and Barbetti 2013, Hua et al 2004)). As such, we believe this continues to be an important record that should be widely discoverable, and ACP is a suitable place for it.
3. Reviewer 2 also asks for a shortening of interpretation that is repeated from previous publications. We understand the reviewer's point of view, but we believe that when reporting on a long record, it is frustrating to the reader to have to refer to previous publications to find interpretation of the long record. We have altered the text to make clear where interpretation has been reported elsewhere and where it is new.

Reviewer 1 Samuel Hammer

Turnbull et al. present a thorough revisit of the entire Wellington atmospheric 14CO2 record. They re-measured archived samples and include new information from tree samples to better investigate known "noisy" periods of original record. Conceivable flagging criteria are formulated and the Wellington record is compared to independent

data sets. Therefore, this manuscript is of upmost scientific interest to the radiocarbon community and I definitely recommend publication in ACP.
In addition to the data review the authors revisit and extend the key findings that the Wellington 14CO2 record provides. For some of the conclusions drawn from the data I would like to ask the authors to reinforce their arguments to overcome my minor concerns.

General comments to the authors:
14C measurements:
Have you investigated if the use of IRMS-13C in the early AMS measurements introduces a bias? Such a potential bias could originate e.g. from a machine immanent fractionation. I assume you have IRMS-13C measurements also for the post-2005 samples. Did you compare the effect of offline and online 13C measurements for the D14C normalization directly? Such an investigation will also quantify the contribution to the scatter which is due to offline 13C analysis in the earlier AMS results.

Yes, of course. We believe that indeed the use of IRMS-13C measurements in the 1995-2005 AMS analyses is the reason for the variability. The very clear reduction in noise from 2005 when online AMS 13C analysis was added is very convincing evidence, and there is ample evidence from many AMS labs that this is likely the explanation. This is discussed in two places (sections 3.3 and 3.5.3). We have expanded the text and pointed the reader to the other section in the discussion.

Smooth curve fit:
Fitting section by section may introduce problems at each overlap of the sections. Wouldn't it be better to use a fit routine which can deal with a changing phase? Pickers et al. mention that STL per se does not require gap filling, only the current implementation of STL does. Pickers et al. also investigate HPspline which would allow for a change in phase. Why didn't you chose this fitting algorithm?
We did consider using other algorithms, particularly STL, since this was used in previous analysis of the Wellington $^{14}CO_2$ record. Pickers et al showed that of the three, HPspline was least able to capture the seasonal cycle of atmospheric records and therefore we did not consider it further.
We agree that the STL technique has the advantage of allowing flexibility in the shape of the seasonal cycle. Instead, this approach assumes that the seasonal cycle and trend vary only slowly over a defined time window. This assumption is problematic for time-series characterized by rapid or abrupt changes, such as radiocarbon. During the bomb peak, the seasonal cycle is dramatically amplified, and it falls off rapidly in the years that follow. When STL is applied to this time-series, the seasonal amplitude is damped during the bomb peak and amplified in the years that follow compared to observations. Likewise, the bomb peak is damped and delayed in the STL estimate of the trend. This can be partially ameliorated by dividing the time-series into sections, but this then leads to the same kinds of overlap issues that you have highlighted as problematic for CCGCRV.

In addition, we found that the gap filling needed in STL was as problematic for this record as the phase problem in CCGCRV – neither is a perfect choice. We take the reviewer's point that STL doesn't necessarily require gap filling, but this would require an entirely new fitting

system that is not currently used in the atmospheric community and would therefore raise a number of questions of its own. Further, the seasonal cycle is quite small after 1979, and the majority of users of this dataset are interested in the annual trend, so overall, we judged that the phasing problem in CCGCRV is less problematic than the gap filling problem of STL. We have made some adjustments in the text to clarify these points, but note that we chose not to explicitly discuss HPspline at all since its limitations have been discussed elsewhere already.

When you investigate the phase change in the 14CO2 signal, you find that the seasonal cycle weakens between 1978 and 1980, and then reverses. Could it be that this timing is related to the change in the fitting sections (1966-1979 and 1980 to1989). The described method for overlap and interpolation between different fits favors the weakening of the seasonal cycle at the section borders if both sections are out of phase. I wonder if you would find the same timing for the phase change if you chose different fitting sections...
In fact, we chose the division at 1979-1980 precisely because the change in seasonal cycle is apparent in the raw observational data at this time period. We tested other divisions into time periods and found that the fitted curve couldn't match the data as well, as diagnosed by the residuals. We have added text to explain our choices.

Hypothesis of reversed seasonal cycles in the early post-bomb era:
The hypothesis behind the changing phase in the seasonal cycle should be backed up by a small (box-) model exercise. This model should include the seasonal cycles of the STE (in NH and SH) and the CEE (cross equator exchange) in the troposphere and the stratosphere. The Mount Pinatubo eruption is a well-studied phenomenon when it comes to stratospheric transport. see e.g. Aquila et al. 2012. They find middle- stratospheric meridional pathways with mixing times of less than a year. The major stratospheric bomb-peak lasted for about 4-5 years (see HASL data compiled in Naegler et al 2006). Can you show in a (box-) model that with those boundary conditions your hypothesis is valid?
Aquila, Valentina, et al. "Dispersion of the volcanic sulfate cloud from a Mount Pinatubo–like eruption." ***Journal of Geophysical Research: Atmospheres*** **117.D6 (2012).**
Please see our general comment to the editor – we spent considerable effort developing a box model, but ultimately demonstrated for ourselves what previous authors had already shown – $^{14}CO_2$ can't be adequately described with a four box model. Instead we have revised our discussion of the seasonal cycles to take on the reviewer comments and utilize previous modelling studies. We have considerably reduced the discussion of the seasonal cycle since the reviewers point out that we don't have sufficient evidence to back it up.

Interpretation of the seasonal cycles since 2005:
I have a couple of questions and comments to the comparison of the Wellington and Cape Grim seasonal cycles:
The comparison to the Cape Grim seasonal cycle is problematic since both mean cycles do not average the same time period. Figure 4b shows that there are obvious large inter-annual variations in the amplitude (phasing?) of the seasonal cycle.
We added a comment that choosing only the period of overlap (2005-2010) gives similar results.

What is the origin of the double maxima in the BHD cycle?
We revised the discussion to make our argument clearer that this is due to transport and STE.

Is the Melbourne influence at Cape Grim detectable in CO2 or CO?
Yes (in the reference provided) – but the in situ and flask data can be screened to remove the local influences, whereas the ^{14}C samples, which reflect the integrated ^{14}C signal over ~two weeks, cannot. Text revised to reflect this point.

Fig 6 does not convince me that BHD is not influenced by anthropogenic emissions. Wellington is in the middle of the "red" area. When reading Pickers et al. they mention that in their data example of the BHD CO2 data they had to gap fil 10% of the data since they deviated from baseline conditions.... To me this indicates some anthropogenic influence at BHD as well.
We expanded the discussion of influences at Baring Head, using the CO_2 observations of Stephens et al (2013) to show that there is a very occasional urban influence, and a more regular terrestrial biosphere influence, but there is no evidence of seasonality in either of these (i.e. they might influence the overall $\Delta^{14}CO_2$ value very slightly, but not the seasonal cycle).

Sure, Melbourne emits 50 times more ffCO2 than Wellington, however the distance between Melbourne and Cape Grim is 340km, whereas it is around 10km between Wellington and the BHD...
We revised the text as in the response above.

If STE is the driving mechanism for the seasonal cycle for the periods 1966 to 1979 and 1980 to 1990, how come that the seasonal cycle post 2005, which is also explained via the STE, is not in phase with the earlier once...
We have removed this argument since the reviewers have pointed out that we don't have sufficient evidence to back it up.

Specific comments:
p.2 l.40 Please state the years when the measurements in Norway and Austria started
Done.

p2. l.44 The term "exchanges" is a bit too general, consider oxidized or something more specific.
Revised. See also reviewer 3 response.

p.2 l.45: Production -> Natural production
Revised.

p2. L 47: perturbations to Δ14CO2 -> perturbations to natural Δ14CO2 levels
We considered this, but on rereading the text, believe that "perturbations to $\Delta^{14}CO_2$" more accurately reflects the point we are trying to convey. In recent years, the fossil fuel perturbation is of great interest, but it is the perturbation relative to the recent atmosphere that we are primarily interested in, not the perturbation relative to natural levels.

p2. L62: Add year to Lopez et al., and add also early attempts of ffCO2 emission estimates like e.g:
Meijer, H. A. J., et al. "Isotopic characterisation of anthropogenic CO 2 emissions using isotopic and radiocarbon analysis."* Physics and Chemistry of the Earth *21.5 (1996): 483-487.
Gamnitzer, U., U. Karstens, B. Kromer, R. E. M. Neubert, H. A. J. Meijer, H. Schroeder, and I. Levin (2006), Carbon monoxide: A quantitative tracer for fossil fuel CO2? J. Geophys. Res., 111, D22302
We added the year to Lopez et al., and added the Meijer paper. There is now a long list of papers that use ^{14}C to understand fossil fuel emissions, and it does not seem appropriate to cite every one of them here. Instead, we tried to list only the key papers describing the method and one from each "scale" of study. We should most definitely include the Meijer et al paper, but the Gamnitzer paper doesn't add much beyond the seminal Levin 2003 paper (at least in this context).

p.2 l77: add citations to the last part of this paragraph
Done.

p4 l128: what do you intend with the term "nominally CO2-free"? Did you process blank NaOH solutions? How much CO2 is in a blank NaOH solution? What is the 14C activity of this blank?
This is in the supplementary material section S3.1 and we added a pointer to that section in the main text.

p4.l131: "large tray" can you state the surface area of that tray?
Added in the supplementary materials.

P4.139: Please add the statement about fractionation (supplement S3.l90-92) to the main text.
Done.

P5 l189 "one" sd? In Fig S2 and the text you state 2 sd?
We have slightly altered the statistical analysis to include a paired sample t test and reworded in both the main text and supplement figure S2 caption to clarify that there is no significant difference between the two methods.

P6 l259 please include a reference to Fig.2 in this subsection
Done.

P8 l316 I don't see the 2005 EN Tandem improvement mentioned in Zondervan et al 2015.... Maybe I overlooked it?
The method for using all three isotopes measured in the AMS is described in Zondervan et al 2015. We moved the reference to the previous sentence to clarify that it applies to the method, not the improvement in precision.

P8 l336 Do the measurements from this period carry a special flag (e.g noisy) in the dataset? Reading the supplement I found that you are already doing this. Maybe make a short note in the main text.
Done.

P8 l353 how does ccgvu handle data gaps and inconsistent sampling frequencies? Since the paper is (at least for me) not freely available it is worth mentioning this shortly in the supplement.
Done.

P9 l362 what is the unit of the cutoff criteria in the frequency domain?
Days. This was a typo.

P9 l363 is the 2 year overlap a good idea? In terms of transition yes, but don't you have now the influence of end-effects in 4 years?
We tested using different overlap periods. Using a shorter overlap causes a nasty end effect jump in the record, and a period that is much longer smooths out the differences too much. We added a comment to justify our choice.

P9 l368 "mean residual difference" do you mean RMS of the residuals
No, this is the mean of the residuals, which are the mean difference between the smooth curve fit and the measured values. We reworded "mean residual difference" to "mean difference", which we think makes this clearer.

P9 l379 state the "n" of the MC
Done.

P9 l382 where are the 95% conf intervals given? In the data set I see only one uncertainty column, please specify in the data-set if this is the 1 sigma error or the 95% conf interval.
We have removed the 95% confidence interval – we had originally included this in the reported dataset, but since it is effectively multiplying the one-sigma uncertainty by two, it seems unnecessary to include it in the final dataset.

P9 l384ff the model simulation are not convincingly not used in the paper. See general comments. Consider skipping the subsection 3.7 and Fig 6.
On lines 479-481 of our original manuscript, we describe the finding of Ziehn et al. [2014] that Cape Grim is influenced by fossil fuel emissions from Melbourne in the wintertime. Ziehn et al show that this seasonal fossil emission influence is primarily driven by seasonal changes in atmospheric transport, rather than seasonality in the fossil fuel emissions. We present the seasonal analysis of our model simulations to demonstrate that Baring Head is not influences by seasonal transport variability. We have clarified the text in this section.

P9 l388 LAU ??
Removed – this was an oversight as the model also generates footprints for the Lauder site (LAU) that are not discussed in this paper.

P10 l403ff include ref to fig. 2
Done.

P10 l442 30 per mil amplitude for the period 1966-1979? I only see such an amplitude once? A mean amplitude of ca. 7 per mil seem more realistic.
We revised the text to clarify that 30‰ is the maximum amplitude and a mean across this period of 6 ‰.

P11 l456 fig 6 -> fig 5??
Removed this figure reference.

P11 l459 "Between 1978 and 1980 the seasonal cycle weakened". This is not really seen in fig 4b.
Unfortunately 1978 to 1980 is a boundary of the fitting sections... since the seasonal cycles for the two sections are opposed and the overlap is linearly interpolate between fits... a weakening can also come from the applied method.
In figure 4b, we added the detrended raw observations and expanded the text.

P11 l460 5 per mil amplitude? Maybe two times in this period... 3 per mil on average
Revised. (we had used peak to peak amplitudes and have revised to use middle-to-peak amplitudes as is more standard).

P11 l467 fig 5 -> fig 4
We intended to refer to figure 5. No change.

P12 l 494 fig5 -> fig 4?
We intended to refer to figure 5. No change.

P12 l497 "records that are indicated in figure 1" -> "records where the sampling locations are indicated in figure 1"
Revised.

P13 l563 Model results from Levin et al. 2010 already suggest the development of a interhemispheric gradient in the same magnitude for the same time... without changing the southern ocean... although they admit that they are not matching the data...
Levin et al. (2010) were the first to suggest the development of an interhemispheric gradient, and we were remiss in our discussion of this. It has been rectified in the revised manuscript. Levin et al. were able to roughly match the observed gradient without changing the Southern Ocean. It is important to note that Levin et al. tuned the terrestrial biosphere component of their model to match the observed global average atmospheric CO_2 and $\Delta^{14}CO_2$. Thus, this paper highlights the fact that a terrestrial process occurring predominantly in the Northern Hemisphere can reproduce the observed gradient, but we do not feel that it proves the gradient was caused by the terrestrial biosphere or rules out a major role for the Southern Ocean.
We have almost entirely re-written this section of the paper to present a more balanced view of the potential processes controlling the gradient. We still feel that a re-organization

of the Southern Ocean is the most likely cause given the supporting evidence from the ocean carbon cycle community. However, we now more clearly acknowledge the previous interpretation of Levin et al. and the fact that we cannot robustly distinguish between a terrestrial and oceanic cause with the existing sparse radiocarbon network.

Table1: include sample no. to NZ/NZA, replace GC with gas counting, change "measurement methods" to "measurement and sampling methods"
We made these changes and added more text to the caption to clarify.

Table2: provide the unit to the 14C differences
Done.

Figure1: provide scales to the google earth pictures, indicate urban areas in the upper map.
Done.

Figure 2. consider vertical grid lines to illustrate the different periods used in the paper.
We tried this but found that it cluttered the graph too much.

Figure 2. Consider indicating graphs with a) and b)
We think (top) and (bottom) are appropriate here since it is quite obvious what is shown.

Figure 2. x-label of graph a) is cropped...
This looks fine in our version. If the problem still appears in the proofs, we will correct it.

Figure 4. Consider indicating graphs with a) and b)
We are happy with using top and bottom.

Figure 4. in a) use the same periods as in the text.
This was a labelling error and has been corrected.

Figure 4. b) consider vertical grid lines to illustrate the different periods
That would be nice, but once we added the detrended observations, adding vertical lines was just too confusing.

Figure 6. Motivate the plot better. Not really used in the paper. Explain the unit.
The unit is now more clearly explained, but we chose to keep this figure for reasons outlined above.

Figure 7. Consider indicating graphs with a) and b)
As before, we are happy with using top and bottom.

Figure 7. Consider usage of open symbols. Especially after 2000 it would be good to see all data.
We tried a number of different ways of presenting this – smaller symbols are hard to see, and open symbols also make it hard to look at. The version we show gave (at least in our opinion) the best presentation of the comparison.

Supplement:
S2.l74 state the surface area of the pyrex tray
Added.

extraction follows -> extraction from 1995 onward follows
Changed.

in total after flagging you have 427 targets, if you split them between the machines you have 397 and 102 .... To me this does not add up? What am I missing?
A mistake on our part – we initially recorded the degrees of freedom in each χ^2_ν calculation rather than the number of targets (degrees of freedom = number of targets – number of unique samples). We have rectified this to give the number of targets.

Please state the main offset for the QC datasets between the two AMS machines.
We expanded this sentence to say that no offset was observed.

S5 l217ff What is RLIMS?
RLIMS is defined in line 36 of the supplement, it is the name of our radiocarbon laboratory database. We added a reminder at this point in the document since the reader might not recall.

S6. L262 Indicate the figure S1 with a) and b). I assume a) is Eastbourne and b) is Baring Head? Correct?
We revised the caption – a is the full record and b is zoomed into the recent time period.

S9 l394 Since you cannot decide between "red" or "green" for the Baring Head tree, how can you than state the excellent agreement? Is it excellent for both red and green? Please include a link to the t-test or the mean difference to reinforce this statement.
See above comment – the different colors indicate where we shifted the ring counts by + or – one year NOT different trees. The bottom graph is simply a zoom of the top one. We revised the caption to make this clearer.

S12.l457 Define "NIK". Why is there only one comparison for NIK and 4 comparisons for BHD?
NIK has been changed to "Eastbourne" (NIK is the short name for the street the trees are located on). Most of the Eastbourne samples are from a single tree, and the comparison between the two trees does not appear to be critical (hence only one comparison). The key comparison is between the BHD and Eastbourne trees, which show no significant differences between $\Delta^{14}CO_2$ at the two locations.

S12.l468 please specify the t-test: I assume you use a dependent t-test for paired samples? Since the applied formulas are easy it might be clearer if you just explicitly state them.
Revised the caption.

what is the mean difference if you use the one year shifted BHD tree (red points in fig S1)?

Shifting the BHD tree one year older gives a mean difference between the BHD and Eastbourne tree rings of 5.6 ± 0.7 ‰ and a paired sample t value of 8. Conversely, shifting one year younger gives a mean difference of -8.4 ± 0.8 ‰ and t of 11. Either shift indicates a poor match and therefore unlikely. We added some text to describe this.

Technical comments:
In the text please use a consistent ordering (e.g. temporally ascending) when citing multiple papers.
Done.

Reviewer 2:
The atmospheric radiocarbon measurements conducted at Wellington are a very important record and the authors' efforts to maintain and evaluate the observations are valuable to the community.
However, there are some major revisions needed before publication of this manuscript. Much of the paper is used on re-reporting trends and gradients that have already been shown in other work. The authors also make unsupported claims about the mecha- nisms driving the interhemispheric gradient and seasonal cycles of D14C.

The paper postulates a sensitivity to Southern Ocean air-sea exchanges that is mis- leading and unsupported. It gives the impression that the Southern Ocean only began influencing the interhemispheric D14C gradient in 2002, whereas the Southern Ocean has always been a primary influence on the interhemispheric D14C gradient, via gross, not net, carbon exchange. Levin et al. 2010 and Randerson 2002 clearly show that the observed trend in the interhemispheric D14C gradient is consistent with a long-term change in the oceanic influence, dominated by the long-term decrease in atmospheric D14C and the change in D14C disequilibrium over the Southern Ocean, which is fur- ther supported by the Graven 2012 papers.
A change in upwelling is interesting to consider as a secondary effect, but the authors do not include quantitative models or estimates of how large the effect could be, nor any specifics on how it influences D14C. Furthermore, the Wellington data from 1995-2005 are shown to have serious issues, which would complicate identification of a signal originating in the early 2000s. And there is no discussion about the period in the 1990s when upwelling was increasing.
We agree with the reviewer that ocean disequilibrium has been important throughout the post-bomb ^{14}C record, and our text was intended to convey that point, and that there is a possibility of a change in upwelling that could change the magnitude of this effect. Based on this and comments from the other reviewers, we have reduced the discussion of the possible change in upwelling, and tried to emphasize that indeed ocean exchange has always been important.

The authors similarly make statements about the influences on the seasonal cycle of D14C at Wellington that aren't well-supported.
Based on this and the other reviewer's comments, we have shortened the seasonal cycle discussion.

The paper should be shortened to minimize the re-reporting of previous observations, reduce repetition, clarify the long-term trend in the Southern Ocean influence on the interhemispheric D14C gradient, and remove unsupported statements. As the main contribution is to revise the Wellington data, i.e. no new modeling or other evidence is given to help interpret the data, the paper might be better suited to a journal like Radiocarbon or Atmospheric Measurement Techniques.
We appreciate the reviewer's view that this work could be well suited to Radiocarbon or AMT. We do believe that the uniqueness of the record, its length and wide use across a large audience makes this worthy of publication in ACP.

Specific Comments.
Section 3.5.3 appears to show major problems in the measurements for the 1995- 2005 period, with large scatter and a high bias. I don't agree that the questionable data should be retained, as the authors have done - "in the absence of better data, we retain both the original and remeasured NaOH sample results in the full record." This conflicts with the aim of the paper to evaluate and refine the previously reported mea- surements and, presumably, to prevent the interpretation of measurement problems as real atmospheric variability.
We believe that it is appropriate to report these results in the observational dataset, rather than simply discarding them from the published record, since we cannot definitively say that they are wrong. We have flagged them clearly in the dataset, and users have the opportunity to use them or discard them. Further, we provide two different fitted curves – one including this data and the other removing it and replacing with Cape Grim data. We have added text to clarify these points.

The code WLG is already used by NOAA for Mt Waliguan, China – perhaps another code would be better.
We have changed the code to BHD, and the actual site (Makara or Baring Head) is still indicated in the data files.

L15 Earliest direct atmospheric
Changed.

L98 Revisiting key findings can be placed in the introduction for brevity.
We believe that the paper is easier to read with the current organization.

L104-108 Unsupported. See above comment.
We removed these sentences from the introduction and shortened the discussion in the results/discussion section.

L234 Please quote a value for precision
Added.

L306 Why would this result in higher D14CO2?
Revised to "This would result in contaminating CO_2 absorbed on the NaOH before the solution was prepared. Since atmospheric $\Delta^{14}CO_2$ is declining, this would result in higher $\Delta^{14}CO_2$ observed in these samples than in the ambient air. "

L378 More detail needed. Where is this used?
We added the following sentence to clarify why this is included: "This is provided for further users of the dataset, and may be particularly helpful when the dataset is used for aging of recent materials."

L384 How do 4-day back trajectories address the seasonal cycle? The panels in the figure all look the same. This is not very useful. A panel should be shown with the differences if there is a difference to highlight.
Ziehn et al. (2014) show that the Cape Grim site is influenced by seasonally coherent changes in the atmospheric transport, such that the site detects fossil fuel emissions from Melbourne in winter but not in other seasons. We show these model simulations precisely to demonstrate that the Baring Head record is not influenced by such seasonal variations in transport. In response to this comment and a similar comment by the first reviewer, we have rewritten and clarified this discussion in the manuscript.

L413 Since 2005 or earlier?
Changed "since" to "after" to clarify.

Section 4.1 seems out of place and repetitive. Should move to introduction and focus on new results here.
We believe that the paper reads more clearly with this discussion here.

L435 Turnbull 2009 only includes simulations from the 2000s, so they do not show the Suess Effect became the dominant driver in the 1990s.
We included references to the two studies that have shown that the Suess Effect is the most important driver after 1990. Levin et al 2010 show this has occurred since 1990, Turnbull 2009 is a second study using an independent model that agrees with the Levin result. We believe it is appropriate to include both references.

L454 Do you mean when mixing with lower-D14C air from the stratosphere was the strongest? Are there Southern Hemisphere stratospheric observations from the bomb period supporting the idea that tropospheric D14C was higher than stratospheric D14C? Are you saying that tropospheric D14C was higher than stratospheric D14C in the Southern Hemisphere until the late 1970s? Bomb 14C would have also entered the SH stratosphere through the tropical tropopause, while at the same time tropo- spheric D14C was declining, so this seems unlikely. Note Northern Hemisphere sites also showed minima in spring in the early bomb period. Levin 2010 simulate recent seasonal influences on D14C and should be cited here. Oceanic influences on the seasonal cycle should also be mentioned.
We have revised this section to remove this discussion.

L468 See Brenninkmeijer, C. A. M., Lowe, D. C., Manning, M. R., Sparks, R. J., & van Velthoven, P. F. J. (1995). The 13C, 14C, and 18O isotopic composition of CO, CH4, and CO2 in the higher southern latitudes lower stratosphere. Journal of Geophysical Research: Atmospheres, 100(D12), 26163-26172. doi:10.1029/95JD02528

Thank you for this reference, but we have removed this discussion and therefore not included it.

L494 This is the time of maximum in the NH so this phasing is unexpected. Is there an explanation for the double-peaked shape of the cycle? This section relies on dismissing the Cape Grim data, which is not entirely convincing. Are other Southern Hemisphere observations relevant here?
There are no other long term records from a similar latitude in the Southern Hemisphere (there are tree ring records, but these clearly cannot resolve seasonal cycles). We expanded this discussion to strengthen our argument. It is worth noting that the seasonal cycle during this period is quite small and the difference between the seasonal cycle in the two records is perhaps 0.5‰.

L517 It would be useful to include a plot of the difference between the Wellington and Cape Grim data.
The two datasets are shown in figure 3 and we added a reference to figure 3 in this sentence.

L521 Delete the word signal. Is it possible to say something more quantitative than "homogeneous"?
In the previous sentence, we say that differences between the two sites are smaller than the measurement uncertainty.

L527 What is the basis for the new estimate of the interhemispheric exchange time? How was this calculated? Without any supporting information this paragraph should be deleted.
We have added further explanation of this calculation. It is surprising that this bomb peak difference has never actually been used to calculate an interhemispheric exchange time before. Although our calculation is simplistic, it agrees nicely with recent, more sophisticated analyses of the exchange time and we think it is worth including.

L544 Need to cite Levin 2010, and Graven 2012
Both are now cited in this paragraph.

L561 Also shown in Randerson 2002 and Levin 2010
We now include the Levin 2010 reference. Randerson 2002 doesn't go beyond 2000 in its data, so it is less relevant here.

L565 This paragraph is misleading. See main comment above.
We have shortened this section considerably.

L575 This is the gross carbon flux not the net carbon flux. Atmospheric D14C has been highly sensitive to Southern Ocean upwelling not only since the 1980s but since the preindustrial period and throughout the bomb peak period – see Randerson 2002 and Levin 2010
We agree with the reviewer, see earlier comments. And have shortened this section considerably.

L593 "Although the changing Southern Ocean carbon sink is the most likely explana- C4 tion," Atmospheric D14C is not directly affected by the Southern Ocean carbon sink. What is the justification for this statement? See main comment above.
See previous comments – shortened this section.

Reviewer 3 J. Miller (Referee)

General comments.

This paper documents and analyzes the longest atmospheric radiocarbon time series from a single site. Obtained near Wellington, New Zealand starting in 1954 and con- tinuing to the present, these data represent a signature time series of carbon cycle science. The authors document the revision and evaluation of the data, which should lead to a significant improvement in its scientific utility. The seasonal cycle and trend are analyzed convincingly, although too much attention is paid to the hypothesis that an increased Southern Ocean CO2 sink can explain the changing Δ14C atmospheric north-south gradient. While it's true that the change in the north-south 14C gradient supports this idea, there is no new analysis of the time series to bolster it. One ad- ditional point is that it would be good to provide the internet location of the data in addition to the static spreadsheet provided. Presumably the ftp site would contain the data set of record including the latest data, flags, and corrections. Nonetheless, this is a strong paper that is entirely appropriate for ACP; it should be published after a few modifications.

The dataset is now available at the WDCCGG and our own websites and we have added a section 7 Data Availability at the end of the text with the links. We are working on also putting the data at CDIAC where much of the global $^{14}CO_2$ data resides, but internal CDIAC issues have slowed this down.

Below, I list some edits and comments by line number.

Specific comments.

L21,22. While Cape Grim air samples may contain anthropogenic signals in winter, air samples have often been collected during times when the wind is not coming from the north.

This is not the case for ^{14}C samples which are integrated over ~2 weeks. We have clarified this key point in the text of our paper.

L44. 'exchanges' is a bit vague. Why not spell it out to say that 14C reacts immediately with O2 to form 14CO, which is subsequently oxidized to 14CO2

Done.

L68-70. This is redundant with text around L44.
The slight repetition seems necessary for the text to be clear. No changes made.

L75. Perhaps strike 'now', and add 'in the two decades following the atm. test ban treaty' at the end of the sentence.
Done.

L77. I don't agree that the additions of fossil fuels became the dominant factor influencing the 14CO2 trend. If fossil fuel CO2 additions are 'dominant' I would think of them being an order of magnitude or so larger than other processes. Presently (and more or less in the 1990s), fossil fuel combustion alone would reduce the atmospheric Δ14C by ~ 10 per mil/yr; cosmogenic production would increase it by 5 per mil/yr; the land- atmosphere and ocean-atmosphere disequilibrium fluxes would be roughly +4 and -4 per mil/yr. It might be reasonable to try and calculate a point at which the negative trend in atmospheric Δ14C was driven more by fossil fuel emissions than by absorp- tion of bomb 14C atoms into the biosphere and oceans. But this would not equate to 'dominant' in my opinion.
This is an important distinction, and we agree with your points. We have changed from "dominant" to "the largest contributor to the $\Delta^{14}CO_2$ trend."

L80. Change 'especial' to 'special'
This is a New Zealand colloquialism. Changed to standard English.

L129. Use 'M' (molar) or 'mol/L'
Done.

L158. 'Faithfully' record Δ14C, but not the 14C content, which is offset by ~ 34 per mil.
Corrected from "^{14}C content" to "Δ^{14}C"

L210. Was testing done do see if the samples could be stored for up to three years before analysis without introducing artifacts.
No such testing has been done, and this is something we will consider for future updates of the record. No changes made to the text.

L216-218. Could using an offline δ13C value produce bias or just add noise? Any tests to examine this?
Yes, this is possible, even likely. We have not done specific tests, but fractionation during sample preparation will almost certainly always go in the same direction. The most likely culprit is incomplete graphitization (in the LG1 graphite system used at this time, reaction completion was not directly measured and we suspect that graphitization was often incomplete), which fractionates to the lighter isotopes and if not diagnosed would result in a higher Δ^{14}C (i.e. goes in the direction of the apparent bias in the data). On the other hand, fractionation in the AMS (most likely in the ion source) is likely to vary in sign through time. We have added explanation in sections 3.3 and 3.5.3 to explain this more clearly.

L227. Considering that the multi-target averaging resulted in differences of up to 5 per mil, I think that this deserves a detailed explanation, at the very least in the supplement.

We agree. An explanation was already given in the supplementary material and we have expanded it slightly and included a note in the main text pointing to the supplement for more information.

L243. S+P's Δ is the same as the presently used Δ14C; their Δ14C is defined differ- ently.
Reworded to clarify.

L255. How was the weighting done? Inverse square of the measurement precision?
Weighted mean as defined by Bevington and Robinson (2003). Sum(x_i*w_i)/sum(w_i), where x_i is the mean of each measurement I and w_i is the weighting, defined $1/w_i$. Since measurement precision does vary, it is appropriate to use a weighted mean rather than a simple mean.

L280. Wondering if 'excursion' is the best word here. Anomaly?
Changed.

L283. As mentioned in comments on L22, Cape Grim sampling can be 'tuned' just for a clean air sector. If the issue is integrated sampling, then I would say that.
Point taken, but rereading this section, the sampling regime at Cape Grim is not germane in this paragraph (although it is relevant elsewhere in the manuscript). No changes made.

L284. Change 'terrestrial' to 'mainland'?
Done.

L303-304. 'preparation was conducted' to 'was prepared'.
Done.

L313. 'or thereafter' to 'and thereafter'
Done. We noticed that one even before the reviewer did.

L325. I don't see the reduction of scatter shown in any plot. It would be useful to show how the reprocessing improved the noise.
The data was not reprocessed, it is that once the change was made, the $\Delta^{14}CO_2$ record immediately becomes less noisy. It is clearly apparent in figure 2b. We added some words in the text to point the reader to the figure.

L351. Change 'ccgvu' to 'ccgcrv' which is the actual name of the curve fitting code.
Done.

L362. Insert 'day' after 80. Good that this important detail was included.
Done.

L395. Add a sentence explaining what a footprint is.
Done.

L403. I think 'roughly "natural"' can be deleted; natural is ambiguous. Maybe 'roughly pre-industrial'?
Done.

L421-422. By 'long-term' to you mean decline since the 1960s? For many in the radiocarbon world, that wouldn't be very long, so maybe define the time period more explicitly. Also, insert 'known' prior to 'long-term trend in. . .'
Revised.

L434. As mentioned earlier, I don't think 'dominant' can be justified.
Changed "dominant" to "largest"

L469. I'm wondering about the value of an untestable hypothesis. What you say sounds plausible, but maybe refer to it as speculation?
We have removed this argument based on reviewer skepticism.

L507. Should Levin et al reference by 2010? 2013 paper appears to deal with Europe.
This is correct in the text – we are referencing the method by which the Jungfraujoch (European) measurements are made.

L527-534. I would like to see the math of how this was calculated, at least in the supplement. Also, one important factor is to know the state of ENSO during the 1963-1965 period, because La Nina, for example, can significantly increase inter-hemispheric exchange. Finally, the SF6 derived value is based purely on surface data, whereas the Δ14C method has a significant upper atmosphere component. It would be good to comment on how the estimates might differ.
This is a very simple calculation – what is the temporal offset between the first maximum of the bomb peak in each hemisphere. We have revised the text to clarify how the calculation was done.

L544 – 596. I felt that the text at the end of the Results and Discussion section focus- ing on the interhemispheric gradient and the Southern Ocean was a bit out of place. The Wellington Δ14C data confirm the gradient observed earlier and extend it in time. However, at present, the two paragraphs (starting at line 565) sound more like a review of the Southern Ocean uptake hypothesis, because there doesn't appear to be any new analysis. If it's not possible to add any new analysis using the Wellington data, I think it would be better to be very concise, essentially saying something like 'our data suggest the S.O sink continues to explain. . . Numerous recent studies using methods x, y and z further support. . . Our data set will be a powerful constraint to understanding the evolution of the gradient in a quantitative model framework. . .'
We have reduced this section to a few sentences. Our intent is to alert readers to the opportunity that Southern hemisphere $^{14}CO_2$ observations give to understanding Southern Ocean carbon cycling.

L571. Change 'natural' to 'mass-dependent'?
Done.

L650. Perhaps acknowledge Scott Lehman and Ingeborg Levin for providing unpub- lished data.
Acknowledgement added. Although we use only published data, they still generously provided the datasets for us to use.

Table 2. WLG is already taken as a site code (for Mt. Waliguan Observatory, China), at least with respect to the WMO GAW program. Wouldn't MAK and BHD work here?
We have changed to use BHD for the overall site code, recognizing that the early part of the record is actually from Makara. However, we want to keep a consistent overall site code so that users are not forced to stitch the two sites together themselves. Another reviewer

raised the same comment.

Figure 2. Can you distinguish the symbols and/or colors for the two versions of the EN-Tandem: i.e. 12,13,14 vs. 13,14, since the results seemed to be significantly different.
Done.

Sixty years of radiocarbon dioxide measurements at Wellington, New Zealand 1954 – 2014

Jocelyn C. Turnbull[1,2,*] Sara E. Mikaloff Fletcher[3], India Ansell[1], Gordon Brailsford[3], Rowena Moss[3], Margaret Norris[1], Kay Steinkamp[3]

[1]GNS Science, Rafter Radiocarbon Laboratory, Lower Hutt, New Zealand
[2]CIRES, University of Colorado at Boulder, Boulder, Colorado, USA
[3]NIWA, Wellington, New Zealand
* contact author: j.turnbull@gns.cri.nz

1. Abstract

We present 60 years of $\Delta^{14}CO_2$ measurements from Wellington, New Zealand (41°S, 175°E). The record has been extended and fully revised. New measurements have been used to evaluate the existing record and to replace original measurements where warranted. This is the earliest direct atmospheric $\Delta^{14}CO_2$ record and records the rise of the ^{14}C "bomb spike", the subsequent decline in $\Delta^{14}CO_2$ as bomb ^{14}C moved throughout the carbon cycle and increasing fossil fuel CO_2 emissions further decreased atmospheric $\Delta^{14}CO_2$. The initially large seasonal cycle in the 1960s reduces in amplitude and eventually reverses in phase, resulting in a small seasonal cycle of about 2 ‰ in the 2000s. The seasonal cycle at Wellington is dominated by the seasonality of cross-tropopause transport, and differs slightly from that at Cape Grim, Australia, which is influenced by anthropogenic sources in winter. $\Delta^{14}CO_2$ at Cape Grim and Wellington show very similar trends, with significant differences only during periods of known measurement uncertainty. In contrast, similar clean air sites in Northern Hemisphere ~~clean air sites~~ show a higher and earlier bomb ^{14}C peak, consistent with a 1.4-year interhemispheric exchange time. From the 1970s until the early 2000s, the Northern and Southern Hemisphere $\Delta^{14}CO_2$ were quite similar, apparently due to the balance of ^{14}C-free fossil fuel CO_2 emissions in the north and ^{14}C-depleted ocean upwelling in the south. The Southern Hemisphere sites show a consistent and marked elevation above the Northern Hemisphere sites since the early 2000s, which is most likely due to reduced upwelling of ^{14}C-depleted and carbon-rich deep waters in the Southern Ocean. This developing $\Delta^{14}CO_2$ interhemispheric gradient is consistent with recent studies that indicate a reinvigorated Southern Ocean carbon sink since the mid-2000s, and suggests that upwelling of deep waters plays an important role in this change.

2. Introduction

Measurements of radiocarbon in atmospheric carbon dioxide ($\Delta^{14}CO_2$) have long been used as a key to understanding the global carbon cycle. The first atmospheric $\Delta^{14}CO_2$ measurements were begun at Wellington, New Zealand in 1954 (Rafter, 1955; Rafter et al., 1959), aiming to better understand carbon exchange processes (Otago Daily Times, 1957). Northern Hemisphere $\Delta^{14}CO_2$ measurements began a few years later in 1962, in Norway (Nydal and Løvseth, 1983) and 1959 in Austria (Levin et al., 1985).

^{14}C is a cosmogenic nuclide produced naturally in the upper atmosphere through neutron spallation, ~~exchanges~~ reacts rapidly to form ^{14}CO and then oxidizes to ~~form~~ $^{14}CO_2$ over a period of 1-2 months, ~~and then~~ after which it moves throughout the global carbon cycle. ~~Production~~ Natural ^{14}C production is roughly balanced by radioactive decay, which mostly occurs in the carbon-rich and slowly overturning ocean carbon reservoir and to a lesser extent in the faster cycling terrestrial carbon reservoir. The perturbations to $\Delta^{14}CO_2$ from atmospheric nuclear weapons testing in the mid-20th century and additions of ^{14}C-free CO_2 from fossil fuel burning have both provided tools to investigate CO_2 sources and sinks.

Penetration of bomb-^{14}C into the oceans has been used to understand ocean carbon uptake processes (Oeschger et al., 1975; Broecker et al., 1985; Key et al., 2004; Naegler et al., 2006; Sweeney et al., 2007). Terrestrial biosphere carbon residence times and exchange processes have also been widely investigated using bomb-^{14}C (e.g. Trumbore et al., 2000; Naegler et al., 2009). Stratospheric residence times, cross-tropopause transport and interhemispheric exchange can also be examined with atmospheric $\Delta^{14}CO_2$ observations (Kjellström et al., 2000; Kanu et al., 2015).

The Suess Effect, the decrease in atmospheric $\Delta^{14}CO_2$ due to the addition of ^{14}C-free fossil fuel CO_2, was first identified in 1955 (Suess, 1955). It has subsequently been refined (Meijer et al., 1996; Levin et al., 2003; Turnbull et al., 2006) and used to investigate fossil fuel CO_2 additions at various scales (e.g. Turnbull et al., 2009a; Djuricin et al., 2010; Miller et al., 2012; Lopez et al., 2013; Turnbull et al., 2015).

The full atmospheric ^{14}C budget has been investigated using long term $\Delta^{14}CO_2$ records in conjunction with atmospheric transport models (Caldiera et al., 1998; Randerson et al., 2002; Naegler et al., 2006; Turnbull et al., 2009b; Levin et al., 2010). These have shown changing controls on $\Delta^{14}CO_2$ through time. Prior to nuclear weapons testing, natural cosmogenic production added ^{14}C in the upper atmosphere, which reacted to CO_2 and moved throughout the atmosphere and the carbon cycle. The short carbon residence time in the biosphere meant that biospheric exchange processes had only a small influence on $\Delta^{14}CO_2$, whereas the ocean exerted a stronger influence due to radioactive decay during its much longer (and temporally varying) turnover time. The addition of bomb ^{14}C in the 1950s and 1960s almost doubled the atmospheric ^{14}C content. This meant that both the ocean and biosphere were ~~now~~ very ^{14}C-poor relative to the atmosphere in the two decades following the atmospheric test ban treaty. As the bomb-^{14}C was distributed throughout the carbon cycle, this impact weakened, and by the 1990s, the additions of

fossil fuel CO_2 became ~~dominant~~the largest contributor to the $\Delta^{14}CO_2$ trend (Randerson et al., 2002; Turnbull et al., 2007; Levin et al., 2010; Graven et al., 2012).

The long-term $\Delta^{14}CO_2$ records have been crucial in all of these findings, and the Wellington $\Delta^{14}CO_2$ record is of ~~e~~special importance, being the oldest direct atmospheric trace gas record, even predating the CO_2 mole fraction record started at Mauna Loa in 1958 (Keeling, 1961; Keeling and Whorf, 2005). It is the only Southern Hemisphere record recording the bomb spike. Several short Southern Hemisphere records do exist (Manning et al., 1990; Meijer et al., 2006; Graven et al., 2012b; Hua and Barbetti, 2013), and some longer records began in the 1980s (Levin et al., 2010). Over the more than 60 years of measurement, there have necessarily been changes in how the Wellington samples are collected and measured. There are no comparable records during the first 30 years of measurement, so that the data quality has not been independently evaluated. Comparison with other records since the mid-1980s has suggested that there may be biases in some parts of the Wellington record (Currie et al., 2011).

Here we present a revised and extended Wellington atmospheric $^{14}CO_2$ record, spanning 60 years from December 1954 to December 2014. We detail the different sampling, preparation and measurement techniques used through the record, compare with new tree ring measurements, discuss revisions to the previously published data and provide a final dataset with accompanying smooth curve fit.

In the results and discussion, we revisit the key findings that the Wellington $^{14}CO_2$ record has provided over the years and expand with new findings based on the most recent part of the record. The most recent publication of this dataset included data to 2005 (Currie et al., 2011) and showed periods of variability and a seasonal cycle at Wellington that differ markedly from the independent Cape Grim, Tasmania $^{14}CO_2$ record at a similar southern latitude (Levin et al., 2010). Here we add complementary new data to investigate these differences, fill gaps and extend the record to near-present. ~~We examine an emerging interhemispheric gradient in $^{14}CO_2$, which supports evidence of a changing Southern Ocean carbon sink. If this emerging gradient is indeed linked to Southern Ocean processes, it suggests that ocean circulation plays a substantive role in the reinvigoration of the Southern Ocean carbon sink.~~

3. Methods

Over 60 years of measurement, a number of different sample collection, preparation, measurement and reporting methods have been used. In this section, we give an overview of the various methods and changes through time, and they are summarized in table 1. Full details of the sampling methods used through time are provided in the supplementary material, compiling methodological information documented in previous reports on the Wellington record (Rafter and Fergusson, 1959; Manning et al., 1990; Currie et al., 2011) along with methods newly applied in this new extension and refinement of the dataset.

3.1. Sampling sites

Samples from 15 December 1954 – 5 June 1987 were collected at Makara (Lowe, 1974), on the south-west coast of the North Island of New Zealand (MAK, 41.25°S, 174.69°E, 300 m asl). Samples since 8 July 1988 have been collected at Baring Head (Brailsford et al., 2012) on the South Coast of the lower North Island and 23 km southeast of Makara (BHD, 41.41°S, 174.87°E, 80 m asl) (figure 1). We also discuss tree ring samples collected from Eastbourne, 12 km north of Baring Head on Wellington Harbour.

3.2. Collection methods

3.2.1. NaOH absorption

The primary collection method is static absorption of CO_2 into nominally CO_2-free 0.5 or 1 M ~~L^{-1}~~ sodium hydroxide (NaOH) solution, which is left exposed to air at the sampling site providing an integrated sample over a period of ~2 weeks (section S3.1; Rafter, 1955). From 1954-1995, ~ 2 L NaOH solution was exposed to air in a large (~450 cm^2 surface area) Pyrex® tray. Since 1995, wide-mouth high-density polyethylene (HDPE) bottles containing ~200 mL NaOH solution were left open inside a Stevenson meteorological screen; the depth of the solution in the bottles remained the same as that in the previously used trays. No significant difference has been observed between the two methods (Currie et al., 2011). A few early (1954-1970) samples were collected using different vessels, air pumped through the NaOH (vs. passive absorption), or NaOH was replaced with barium hydroxide (Rafter, 1955; Manning et al., 1990). CO_2 is extracted from the NaOH solution by acidification followed by cryogenic distillation (Rafter and Fergusson, 1959; Currie et al., 2011). Static NaOH absorption necessarily fractionates relative to CO_2 in the atmosphere. Typical $\delta^{13}C$ values are -15 to -25 ‰ for these samples, and this is corrected for in the data analysis.

3.2.2. Whole air flasks

In this study, we use whole air flask samples collected at Baring Head to supplement and/or replace NaOH samples. Flasks of whole air are collected by flushing ambient air through the flask for several minutes then filled to slightly over ambient pressure. Most flasks were collected during southerly, clean air conditions (Stephens et al., 2013). CO_2 is extracted cryogenically (Turnbull et al., 2015). For whole air samples collected from 1984-1993, the extracted CO_2 was archived until 2012. We evaluated the quality of this archived CO_2 using two methods. Tubes with major leakage were readily detected by air present in the tube and were discarded. $\delta^{13}C$ from all the remaining samples was in agreement with $\delta^{13}C$ measured from separate flasks collected at Baring Head and measured for $\delta^{13}C$ by Scripps Institution of Oceanography at close to the time of collection (http://scrippsco2.ucsd.edu/data/nzd). Whole air samples collected since 2013 are analyzed for $\delta^{13}C$ and other trace gases and isotopes at NIWA (Ferretti et al., 2000) and for the $^{14}CO_2$ measurement, CO_2 is extracted from whole air at Rafter Radiocarbon Laboratory (Turnbull et al., 2015).

3.2.3. Tree rings

When trees photosynthesize, they faithfully record the $\Delta^{14}C$ ~~^{14}C~~ ~~content~~ of ambient CO_2 in their cellulose, the structural component of wood. Annual tree rings therefore provide

a summertime (approximately September – April in the Southern Hemisphere) daytime average $\Delta^{14}CO_2$. Photosynthetic uptake varies during the daylight hours depending on factors including growth period, sunlight, and temperature (Bozhinova et al., 2013), resulting in a somewhat different effective sampling pattern than the 1-2 week NaOH solution collections. We show in section 3.5.1. that at the Wellington location this difference is negligible. Note that we assign the mean age of each ring as January 1 of the year in which growth finished (i.e. the mean age of a ring growing from September – April), whereas dendrochronologists assign the "ring year" ~~is~~ as the year in which ring growth started (i.e. the previous year).

We collected cores from three trees close to the Baring Head site. A pine (*Pinus radiata*) located 10 m from the Baring Head sampling station (figure 1) yielded rings back to 1986 (Norris, 2015). A longer record was obtained from two New Zealand kauri (*Agathis australis*) specimens planted in 1919 and 1920, located 20 m from one another in Eastbourne, 12 km from Baring Head (figure 1). Kauri is a long-lived high-density ~~hardwood~~ softwood species that has been widely used in dendrochronology and radiocarbon calibration studies (e.g. Hogg et al., 2013).

Annual rings were counted from each core. Shifting the Eastbourne record by one year in either direction moves the ^{14}C bomb spike maximum out of phase with the NaOH-based Wellington $\Delta^{14}CO_2$ record (supplementary figure S1), confirming that the ring counts are correct. For the Baring Head pine, rings go back to only 1986, and we verify them by comparing with the Eastbourne record. They show an insignificant mean difference of -0.4 ± 0.8 ‰ (supplementary figure S1).

In practice, it is difficult to ensure that one annual ring is sampled without losing any material from that ring, and no wood from surrounding rings is included. To evaluate the potential bias from this source, we measured replicate samples from different cores from the same tree (Baring Head) or two different trees (Eastbourne~~, 12 km north of Baring Head~~). For samples collected since 1985, all these replicates ~~agree within one standard deviation~~are consistent within their assigned uncertainties (supplementary figure S2). However, for three replicates from Eastbourne in 1963, 1965 and 1971, we see large differences of 9.2, 44.5 and 4.9 ‰, which we attribute to small differences in sampling of the rings that were magnified by the rapid change in $\Delta^{14}C$ of up to 200 ‰ yr^{-1} during this period. Thus, the tree ring $\Delta^{14}C$ values during this period should be treated with caution.

Cellulose was isolated from whole tree rings by first removing labile organics with solvent washes, then oxidation to isolate the cellulose from other materials (Norris, 2015; Hua et al., 2000). The cellulose was combusted and the CO_2 purified following standard methods in the Rafter Radiocarbon Laboratory (Baisden et al., 2013).

3.3. ^{14}C measurement

Static NaOH samples were measured by conventional decay counting on the CO_2 gas from 1954 – 1995 (Manning et al., 1990; Currie et al., 2011) and these are identified by their unique "NZ" numbers. All measurements made since 1995, including recent measurements of flask samples collected in the 1980s and 1990s, were reduced to

graphite, measured by accelerator mass spectrometry (AMS), and are identified by their unique "NZA" numbers. The LG1 graphitization system was used from 1995 to 2011 (NZA < 50,000) (Lowe et al., 1987), and replaced with the RG20 graphite system in 2011 (NZA > 50,000) (Turnbull et al., 2015). Samples measured by AMS were stored for up to three years between sample collection and extraction/graphitization/measurement.

For samples collected from 1995 to 2010, an EN Tandem AMS was used for measurement (NZA < 35,000, Zondervan and Sparks, 1996). Until 2005 (NZA <30,000, including all previously reported Wellington $^{14}CO_2$ data), only ^{13}C and ^{14}C were measured on the EN Tandem system, so the normalization correction for isotopic fractionation (Stuiver and Polach, 1977) was performed using an offline isotope ratio mass spectrometer $\delta^{13}C$ value. The data reported from 2005 onwards (NZA > 30,000) show a reduction in scatter reflecting the addition of online ^{12}C measurement in the EN Tandem system in 2005. This allows direct online correction for isotopic fractionation ~~that~~ that may occur during sample preparation and in the ~~accelerator~~ AMS ~~itself~~ system (Zondervan et al., 2015), and results in improved long-term repeatability. Fractionation in the AMS system may vary in sign depending on the particular conditions, but incomplete graphitization biases the graphite towards lighter isotopes, which, if undiagnosed, will bias $\Delta^{14}C$ high. The LG1 graphitisation system used during this period did not directly evaluate whether graphitization was complete, so it is possible or even likely that there was a high bias in the 1995 – 2005 measurements. This is further discussed in section 3.5.3.

For all EN Tandem samples, a single large aliquot of extracted CO_2 was split into four separately graphitized and measured targets and the results of all four were averaged. We have revisited the multi-target averaging, applying a consistent criterion to exclude outliers and using a weighted mean of the retained measurements (supplementary material). This results in differences of up to 5 ‰ relative to the values reported by Currie et al. (2011) and is discussed in more detail in the supplementary material.

In 2010, the EN Tandem was replaced with a National Electrostatics Corporation AMS, dubbed XCAMS (NZA > 34,000). XCAMS measures all three carbon isotopes, such that the normalization correction is performed using the AMS measured ^{13}C values (Zondervan et al., 2015). XCAMS measurements are made on single graphite targets measured to high precision of typically 1.8 ‰ (Turnbull et al., 2015).

3.4. Results format

NaOH samples are collected over a period of typically two weeks, and sometimes much longer. We report the date of collection as the average of the start and end dates. In cases where the end date was not recorded, we use the start date. For a few samples, the sampling dates were not recorded or are ambiguous, and those results have been excluded from the reported dataset.

Results are reported here as $F^{14}C$ (Reimer et al., 2004) and $\Delta^{14}C$~~^{14}C~~ (Turnbull et al., 2007). $F^{14}C$ is corrected for isotopic fractionation and blank corrected. We calculated $F^{14}C$ from the original measurement data recorded in our databases, and updated a

handful of records where transcription errors were found. $\Delta^{14}C$ is derived from $F^{14}C$, and corrected for radioactive decay since the time of collection; this is slightly different from $\Delta^{14}C$ as defined by Stuiver and Polach (1977) that is corrected to the date of measurement. $\Delta^{14}C$ has been recalculated using the date of collection for all results, resulting in changes of a few tenths of permil in most $\Delta^{14}C$ values relative to those reported by Currie et al. (2011) and Manning et al. (1990). Uncertainties are reported based on the counting statistical uncertainty and for AMS measurements we add an additional error term, determined from the long-term repeatability of secondary standard materials (Turnbull et al., 2015). Samples for which changes have been made relative to the previously published results are indicated by the quality flag provided in the supplementary dataset. Where more than one measurement was made for a given date, we report the weighted mean (Bevington and Robinson, 2003) of all measurements.

3.5. Data validation

3.5.1. Tree ring comparison

Over the more than 60 years of the Wellington $\Delta^{14}CO_2$ record, there have necessarily been many changes in methodology, and the tree rings provide a way to validate the full record, albeit with lower resolution. Due to the possible sampling biases in the tree rings (section 3.2.3.), we do not include them in the final updated record, but use them to validate the existing measurements.

During the rapid $\Delta^{14}CO_2$ change in the early 1960s, there are some differences between the kauri tree ring and Wellington $\Delta^{14}CO_2$ records (Figure 2). The 1963 and 1964 tree ring samples are slightly lower than the concurrent $\Delta^{14}CO_2$ samples. The peak $\Delta^{14}CO_2$ measurement in the tree rings is 30 ‰ lower than the smoothed $\Delta^{14}CO_2$ record, and 100‰ lower than the two highest $\Delta^{14}CO_2$ measurements in 1965. These differences are likely due to small errors in sampling of the rings, which will be most apparent during periods of rapid change.

Prior to 1960 and from the peak of the bomb spike in 1965 until 1990, there is remarkable agreement between the tree rings and Wellington $\Delta^{14}CO_2$ record, with ~~the wiggles~~ the variability ~~in the record~~ replicated in both records. And since 2005, there is excellent agreement across all the different records. Some differences are observed in 1990-1993 and 1995-2005, which we discuss in the following sections.

3.5.2. 1990-1993 ~~excursion~~anomaly

An ~~excursion~~ anomaly in the gas counting measurements between 1990 and 1993 has previously been noted (figures 2, 3) as a deviation from the Cape Grim $\Delta^{14}CO_2$ record (Levin et al., 2010) during the same period. Cape Grim is at similar latitude, and observes a mixture of air from the mid-latitude Southern Ocean sector and ~~terrestrial~~ mainland Australia (Ziehn et al., 2014; Law et al., 2010). The Wellington and Cape Grim records overlap during almost all other periods (figure 3).

We use archived CO_2 from flask samples to evaluate this period of deviation. First, the recent flask samples collected since 2013 (n=12) agree very well with the NaOH static samples from the same period (figure 2), indicating that despite the difference in

sampling period for the two methods, flask samples reflect the $\Delta^{14}CO_2$ observed in the longer-term NaOH static samples. We then selected a subset of archived 1984 - 1992 extracted CO_2 samples for measurement, mostly from Southerly wind conditions, but including a few from other wind conditions. These flask $\Delta^{14}CO_2$ measurements do not exhibit the ~~excursion~~ anomaly seen in the NaOH static samples (figure 2), implying that the deviation observed in the original NaOH static samples may be a consequence of sampling, storage or measurement errors. Annual tree rings from both the kauri and pine follow the flask measurements for this period (figure 2), confirming that the NaOH static samples are anomalous.

The 1990-1993 period was characterized by major changes in New Zealand science, both in the organizational structure and personnel. Although we are unable to exactly reconstruct events at that time, we hypothesize that the NaOH solution was ~~preparation was conducted~~prepared slightly differently, perhaps omitting the barium chloride precipitation step for these samples. This would result in contaminating CO_2 absorbed on the NaOH before the solution was prepared. Since atmospheric $\Delta^{14}CO_2$ is declining, this ~~which~~ would result in higher $\Delta^{14}CO_2$ observed in these samples than in the ambient air. Another possibility is that there were known issues with the background contamination in the proportional counters during this period that could result in a high bias $\Delta^{14}CO_2$. In any case, these values are anomalous and we remove the original NaOH static sample measurements between 1990 and 1993 and replace them with the new flask measurements for the same period.

3.5.3. 1995-2005 variability

As already discussed in section 3.3, the measurement method was changed from gas counting to AMS for samples collected in 1995 ~~or~~ and thereafter. During the first ten years of AMS measurements, the record is much noisier than during any other period (figure 2). Until 2005, offline $\delta^{13}C$ measurements on the evolved CO_2 were used in the normalization correction. In 2005, online ^{12}C measurement was added to the AMS system, allowing online AMS measurement of the $\delta^{13}C$ value and accounting for any fractionation during sample preparation and AMS measurement (Zondervan et al., 2015; see also section 3.3). This s~~, s~~ubstantially ~~improving~~ improved the measurement accuracy and the noise in the $\Delta^{14}CO_2$ record immediately reduced as can be seen in the lower panel of figure 2. Therefore we suspect that the variability and apparent high bias in the 1995-2005 period of the $\Delta^{14}CO_2$ record is due to measurement uncertainty and bias rather than atmospheric variability.

The remaining NaOH solution for all samples collected since 1995 has been archived, and typically only every second sample collected was measured, with the remainder archived without ~~sampling~~extraction. In 2011-2016, we revisited the 1995-2005 period, remeasuring some samples that had previously been measured and some that had never been measured for a total of 52 new analyses.

The new measurements for~~on~~ this time period do show reduced scatter over the original analyses, particularly for the period from 1998-2001 where the original analyses appear anomalously low and in 2002-2003 when the original analyses appear anomalously high.

Yet there remain a number of both low and high outliers in the new measurements. These are present in both the samples that were remeasured and in those for which this was the first ~~sample~~ extraction ~~from~~ of the ~~bottle~~sample. This suggests that a subset of the archived sample bottles were either contaminated at the time of collection, or that some bottles were insufficiently sealed, causing contamination with more recent CO_2 during storage. Comparison with the tree ring measurements and with the Cape Grim record (Levin et al., 2010) suggest that the measurements during this period may, on average, be biased high as well as having additional scatter (figure 3). Nonetheless, in the absence of better data, we retain both the original and remeasured NaOH sample results in the full Wellington record, with a special flag to allow users to easily remove the questionable results if they prefer. We also provide a smoothed fit that excludes these data (section 3.6).~~record.~~

3.6. Smooth curve fit

In addition to the raw measured $\Delta^{14}CO_2$ values, we calculate a smooth curve fit and deseasonalized trend from the Wellington $\Delta^{14}C$ and $F^{14}C$ datasets. The deseasonalized trend may be more useful than the raw data for aging of recent materials (e.g. Reimer et al., 2004; Hua et al., 2013). Acknowledging that the 1995-2005 period is variable and possibly biased in the Wellington record, we also provide in the supplementary material an alternative mid-latitude Southern Hemisphere smooth curve fit and deseasonalized trend in which the Wellington data for 1995-2005 has been removed and replaced with the Cape Grim data for that period (Levin et al., 2010).

Curvefitting is particularly challenging for the $\Delta^{14}CO_2$ record, since (a) there are data gaps and inconsistent sampling frequency, (b) the growth rate and trend vary dramatically and (c) the seasonal cycle changes both in magnitude and phase (section 4.2). We chose to use the ~~ccgvu~~CCGCRV fitting procedure (Thoning et al., 1989), which uses fast Fourier transform and low-pass filtering techniques to obtain a smoothed seasonal cycle and long term trend from atmospheric data. This technique can readily handle the data gaps, ~~and~~ inconsistent sampling frequency~~-~~, and rapid changes in the seasonal cycle and trend. ~~in our record~~T~~, whereas the other widely used fitting procedure, seasonal trend decomposition using locally weighted scatter plot smoothing (STL) requires gap-filling for our dataset (e.g.(Cleveland et al. 1995; Pickers et al., 2015).) assumes that the seasonal cycle and trend change only gradually over a specified time period. This assumption is problematic for the radiocarbon timeseries, due to the rapid changes in the trend and seasonal amplitude during and following the bomb spike.~~

~~requires gap-filling for our dataset~~ (e.g. ~~Pickers et al., 2015). However, ccgvu~~CCGCRV assigns a single set of harmonic terms across the full time period, which is inappropriate in this case of large variation in the seasonal cycle. Thus, we separate the record into five time periods: 1954-1965, 1966-1979, 1980-1989, 1990-2004, 2005-2014. These divisions were chosen ~~as periods when~~ based on major changes in the raw observational growth rate, seasonal cycle and data quality ~~change: 1954-1965, 1966-1979, 1980-1989, 1990-2004, 2005-2014~~. The peak of the bomb spike in the Southern Hemisphere (1965) results in a very large change in seasonality that makes an obvious cutoff point. There is an obvious change in seasonality in the raw observations in 1979 – 1980. The 1990 to

2004 period was grouped to include the time when flask measurements have ~~replaced~~supplemented original NaOH measurements, and the 1995 – 2004 period with noisy data.

The other widely used fitting procedure, seasonal trend decomposition using locally weighted scatter plot smoothing (STL, Cleveland et al. 1990; Pickers et al., 2015) assumes that the seasonal cycle and trend change only gradually over a specified time period. This assumption is problematic for the $\Delta^{14}CO_2$ time-series, due to the rapid changes in the trend and seasonal amplitude during and following the bomb spike. Using this method would necessitate both gap-filling the record and dividing the record into time periods (as we have done for CCGCRV), giving no advantage over CCGCRV.

For each time period, we use ~~ccgvu~~CCGCRV with one linear and two harmonic terms and fit residuals are added back using a low-pass filter with an 80 day cutoff in the frequency domain. At each transition, we overlapped a two-year period and linearly interpolated the two fits across that two year period to smooth the transitions caused by end effects. We tested different overlap periods, and found that two years was optimal to minimize end effects and retain the benefit of separating the time periods. The deseasonalized trend was determined from the full dataset rather than the five time periods, as it does not include the seasonality and produces the same result in either case.

We tested other time period divisions, and our chosen time divisions have the lowest mean residual difference from the measured $\Delta^{14}CO_2$, indicating the best fit to the data (we tested only periods of >10 years since it is difficult to draw conclusions about seasonal cycles from shorter periods when the seasonal cycle amplitude is small relative to the measurement uncertainty). The mean ~~residual~~ difference between the fitted curve and the measured $\Delta^{14}CO_2$ values is 3.8 ‰, consistent with the typical measurement uncertainty for the full dataset. Further, the residuals are highest for the early period (1954-1970) at 6 ‰, consistent with the larger measurement errors at that time of ~6 ‰. The residuals improve as the measurement errors reduce, such that since 2005, the mean residual is 2 ‰, consistent with the reported 2 ‰ uncertainties. The exception is the 1995- 2005 period where the mean residual difference of 5 ‰ is substantially higher than the mean reported uncertainty of 2.5 ‰, reflecting the apparent larger scatter during this period as discussed in section 3.5.3.

The one-sigma uncertainty on the smoothed curve and deseasonalized trend were determined using a Monte Carlo technique (n=100). Each data point was perturbed by a random normal error based on the reported uncertainty of that data point, such that the standard deviation of all perturbations would equal the reported uncertainty to derive the one-sigma uncertainty ~~and 95% confidence interval~~ for the smooth curve. This is provided for further users of the dataset, and may be particularly helpful when the dataset is used for aging of recent materials.

3.7 Atmospheric Model Simulations

Simulations from the Numerical Atmospheric dispersion Modelling Environment (NAME) III Lagrangian dispersion model (Jones et al., 2007) were used to interpret

seasonal variability in the dataset. The NAME model is run backwards in time to analyse the history of the air traveling towards BHD ~~and LAU~~ over the preceding 4 days. For each day of the simulation period, 10,000 particles were released during two time windows in the afternoon; 13:00-14:00 and 15:00-16:00. NAME was driven by meteorological output from the New Zealand Limited Area Model-12 (NZLAM-12), a local configuration of the UK Met Office Unified Model (Davies et al., 2007.) NZLAM has a horizontal resolution of ~12 km, with 70 vertical levels ranging from the earth's surface to 80km. These simulations have been described in more detail by Steinkamp et al. (2016). When these daily simulations are integrated over an extended period of time, they comprise a 'footprint' of the catchment area observed by the site over that period. ~~The average footprints presented here were computed by summing the footprints for every day and release period in 2011-2013 and normalizing them such that the domain integral equals one.~~

Comment [JCT1]: Sara – reviewer request was to "add a sentence to explain what footprints are" – can you improve on what I've added?

4. Results and Discussion

4.1. Variability in the Wellington record through time

The Wellington $\Delta^{14}CO_2$ record begins in December 1954, at a roughly ~~"natural" pre-bomb~~pre-industrial $\Delta^{14}CO_2$ level of -20 ‰ (figure 2). From 1955, $\Delta^{14}CO_2$ increased rapidly, near doubling to 700 ‰ in 1965 at Wellington, due to the production of ^{14}C during atmospheric nuclear weapons tests. Nuclear tests in the early 1950s contributed to the rise, then a hiatus in testing in the late 1950s led to a plateau in Wellington $\Delta^{14}CO_2$ before a series of very large atmospheric tests in the early 1960s led to further increases (Rafter and Ferguson, 1959; Manning et al., 1990).

Most atmospheric nuclear weapons testing ceased in 1963, and the Wellington $\Delta^{14}CO_2$ record peaks in 1965 then begins to decline, at first rapidly at -30 ‰ yr^{-1} in the 1970s and gradually slowing to -5 ‰ yr^{-1} ~~since~~ after 2005. The initial rapid decline has been attributed primarily to the uptake of the excess radiocarbon into the oceans, and to a lesser extent, uptake into the terrestrial biosphere (Naegler et al 2006; Randerson et al., 2002; Manning et al., 1990; Stuiver and Quay 1981). The short residence time of carbon in the biosphere means that from the 1980s, the terrestrial biosphere changed from a ^{14}C sink to a ^{14}C source as the bomb pulse was re-released (Randerson et al., 2002; Levin et al., 2010).

Natural cosmogenic production of ^{14}C damps the rate of ~~long term~~ decline since the bomb peak ~~, increasing $\Delta^{14}CO_2$~~ by ~5 ‰ yr^{-1} in $\Delta^{14}CO_2$; this may vary with the solar cycle, but there is no known long-term trend in this component of the signal (Turnbull et al., 2009; Naegler et al., 2006). There is also a small positive contribution from the nuclear industry which emits ^{14}C to the atmosphere, and this has increased from zero in the 1950s to 0.5 – 1 ‰ yr^{-1} in the last decade (Turnbull et al., 2009b; Levin et al., 2010; Graven and Gruber, 2011).

The Suess Effect, the decrease in atmospheric $\Delta^{14}CO_2$ due to the addition of ^{14}C-free fossil fuel CO_2 to the atmosphere (Suess, 1955; Tans, 1979; Levin et al., 2003), was first recognized in 1955 and has played a role throughout the record. Although the magnitude of fossil fuel CO_2 emissions has grown through time, when convolved with the declining atmospheric $\Delta^{14}CO_2$ history, the impact on $\Delta^{14}CO_2$ has stayed roughly constant at -10 ‰ yr^{-1} since the 1970s (Randerson et al., 2002; Levin et al., 2010). Since the 1990s, the Suess Effect has been the ~~dominant~~ largest driver of the ongoing negative growth rate (Turnbull et al., 2009b; Levin et al., 2010).

4.2. Seasonal variability in the Wellington record

We determine the changing seasonal cycle from smooth curve fits to five separate periods of the record (1954-1965, 1966-1979, 1980-1989, 1990-2004, 2005-2014, figure 4 top panel). This subdivision is necessary to allow the seasonal cycle to vary through time since the ~~ccgvu~~CCGCRV curve fitting routine assigns a single set of harmonics to the time period fitted (see section 3.6). We also created detrended $\Delta^{14}CO_2$ values by subtracting the deseasonalised trend from the observations. Comparison with the

detrended fitted seasonal cycle determined from the smooth curve fits (figure 4 bottom panel) shows that the smooth curve fit, as might be expected, does not capture the largest deviations from the trend seen in the observations, but represents the changing seasonal cycle quite well.

The 1966-1979 period shows a strong seasonal cycle (figure 4~~) of about 30‰ amplitude~~) with a consistent phase and an amplitude that varies from a maximum ~~of~~ in 1966 of 30‰ gradually declining to 3 ‰ in 1979, ~~and~~ with a mean amplitude of about 6 ‰. This ~~,~~ ~~which~~ is primarily attributed to seasonally varying stratosphere – troposphere exchange bringing bomb ^{14}C into the troposphere (Manning et al., 1990; Randerson et al., 2002). Manning et al. (1990) were unable to simulate the correct phasing of the seasonal cycle, apparently because their model distributed bomb ^{14}C production throughout both Northern and Southern stratosphere. In fact, the majority of the bomb ^{14}C was produced in the Northern Hemisphere stratosphere (Enting et al., 1982). Randerson et al (2002) were able to match the amplitude of the Wellington seasonal cycle during this time period, although their model was out of phase with the observations by about 1.5 months. They attribute the seasonal cycle during this period mostly to the seasonality in Northern Hemisphere stratosphere – troposphere exchange with a phase lag caused by cross-equator exchange into the Southern Hemisphere. The seasonal cycle kept the same phase but gradually decreased in amplitude until the late 1970s, attributed to the declining disequilibrium between the stratosphere and troposphere as the bomb ^{14}C moved throughout the carbon reservoirs.

~~and we show schematically in figure 5 why this causes the opposite seasonal phase. Most transport across the equator occurs in the troposphere, so that the Southern Hemisphere stratosphere would have had a lower $\Delta^{14}CO_2$ than the Southern Hemisphere troposphere during the early post-bomb period (figure 6). Since maximum cross-tropopause exchange occurs in the spring (Olsen et al., 2003), this resulted in a minimum in $\Delta^{14}CO_2$ at Wellington in the austral spring (August) when bomb ^{14}C moved most rapidly into the stratosphere. The seasonal cycle kept the same phase but gradually decreased in amplitude until the late 1970s, attributed to the declining disequilibrium between the stratosphere and troposphere as the bomb ^{14}C moved throughout the carbon reservoirs.~~

Between 1978 and 1980 the seasonal cycle weakened, and then reversed during the 1980s, with a maximum in winter (June – August) and amplitude of about ~~5~~ 2 ‰. The detrended observations show that this change in phase is not an artifact of the fitting method (bottom panel of figure 4). This result is comparable to that obtained by Manning et al. (1990) and Currie et al. (2011), who both used a seasonal trend loess (STL) procedure to determine the seasonal cycle from the same data. This is consistent with a change in sign of the terrestrial biosphere contribution as the bomb ^{14}C pulse began to return to the atmosphere from the biosphere (Randerson et al., 2002).~~We hypothesize that as tropospheric $\Delta^{14}CO_2$ declined, and continued natural production of ^{14}C occurred in the stratosphere, the Southern Hemisphere stratosphere eventually became enriched in ^{14}C relative to the Southern Hemisphere troposphere, so that consistent seasonally varying exchange processes resulted in a change in sign of cross-tropopause $\Delta^{14}CO_2$ exchange in the late 1970s (figure 5). To the best of our knowledge, no Southern Hemisphere~~

~~stratosphere $\Delta^{14}CO_2$ measurements have been made since the mid-1970s, so there is no direct evidence for this hypothesis.~~

The Wellington $\Delta^{14}CO_2$ seasonal cycle declined in the 1990s, and the larger variability in the observations between 1995 and 2005 makes it difficult to discern a seasonal cycle during that period. Since 2005, the more precise measurements allow us to detect a small seasonal cycle with amplitude of about 2 ‰ (figure 4). We compare the seasonal cycle at Wellington from 2005 – 2015 with ~~Measurements from~~the seasonal cycle at Cape Grim, Australia from 1995-2010. There is no significant difference in the seasonal cycle at either site if we select only the overlapping time period of 2005-2010. Both sites show a similar magnitude seasonal cycle ~~to that at Wellington from 2005 – 2015~~during this ~~time~~ period, and Cape Grim shows a maximum in March – April that has been attributed primarily to the seasonality of atmospheric transport of Northern Hemisphere fossil fuel emissions to the Southern ~~p~~troposphere (Levin et al., 2010). This maxima at Cape Grim coincides with a seasonal maximum in the Wellington record. ~~(Levin et al., 2010).~~ However, Wellington $\Delta^{14}CO_2$ exhibits a second maximum in the austral spring (October) that is not apparent at Cape Grim.

Recent work has shown that during the winter, the Cape Grim station is influenced by air coming off the Australian mainland including the city of Melbourne (Ziehn et al., 2014), which would act to reduce $\Delta^{14}CO_2$ at Cape Grim relative to Southern Ocean clean air. This shift is shown to be the result of seasonal variations in atmospheric transport. ~~In contrast, the Baring Head location near Wellington is not significantly influenced by urban regions in any season (figure 6).~~ The two-week integrated sampling used for $\Delta^{14}CO_2$ at both Cape Grim and Baring Head means that in contrast to other species, $\Delta^{14}CO_2$ measurements cannot be screened to remove these pollution events.

In contrast, the Baring Head location near Wellington does not show significant seasonal variation in atmospheric transport (figure 5) and Baring Head is less likely than Cape Grim to be influenced by anthropogenic emissions in any season. Air is typically from the ocean, and the local geography means that ~~Air is generally from the ocean, and local topography means that the wind is almost always either from the north or south, so that In contrast, the Baring Head location near Wellington is not significantly influenced by urban regions in any season (figure 6). Air is typically from the ocean, and the local geography means that~~ the urban emission plume from Wellington and its northern suburbs of Lower Hutt very rarely passes over Baring Head (figure 1) and the typically high wind speeds further reduce the influence of the local urban area (Stephens et al., 2013). During the austral autumn, there is some land influence from the Christchurch region in the South Island, but emissions from Christchurch are much smaller than the Melbourne emissions influencing Cape Grim: State of Victoria fossil fuel CO_2 emissions for 2013 were 23 MtC whereas Wellington and Christchurch each emitted 0.4 MtC of fossil fuel CO_2 in 201~~2~~3~~/13~~ (Boden et al., 2012; AECOM, 2016; Australian Government, 2016).

Although broad-scale flow from the west is common (figure 5), the local topography means that local air flow is almost always either southerly or northerly (Stephens et al.,

2013), but during rare (<5% of the time) westerly wind events, fossil fuel emissions from Wellington do appear to cause enhancements of up to 2 ppm in CO_2 (Stephens et al., 2013), which would decrease $\Delta^{14}CO_2$ by ~1 ‰ during such an event. Yet there is no evidence of seasonality in the infrequent westerly events ~~(Figure 6)~~. Northerly conditions bring a terrestrial biosphere influence that elevates CO_2 by about 1 ppm (Stephens et al., 2013), which could result in a maximum increase in $\Delta^{14}CO_2$ of ~0.2‰ relative to background conditions, but there is no evidence that this influence is seasonally variable either. Thus, although there are some local influences on the Baring Head $\Delta^{14}CO_2$, none of these appear to be seasonally dependent and instead, t~~T~~he observed Baring Head $\Delta^{14}CO_2$ maximum in spring in the recent part of the record ~~can~~ may be explained by the seasonal maximum in cross-tropopause exchange bringing ^{14}C-enriched air at this time of year ~~(figure 5)~~.

4.3. Comparison with other atmospheric $\Delta^{14}CO_2$ records

We compare the Wellington $\Delta^{14}CO_2$ record with several other $\Delta^{14}CO_2$ records, located as~~that are~~ indicated in figure 1. First, we compare with measurements from Cape Grim, Australia (CGO, 40.68°S, 144.68°E, 94 m asl). Cape Grim is at similar latitude to Wellington and also frequently receives air from the Southern Ocean (Levin et al., 2010). Samples are collected by a similar method to the Wellington record using NaOH absorption and are measured by gas counting to ~2 ‰ precision. Next, we compare with mid-latitude high-altitude clean air sites in the Northern Hemisphere. The Vermunt, Austria (VER, 47.07°N, 9.57°E, 1800 m asl) record began in 1958, only a few years after the Wellington record began, and in the 1980s the site was moved to Jungfraujoch, Switzerland (JFJ, 46.55°N, 7.98°E, 3450 m asl); these measurements are made in the same manner and by the same laboratory as the Cape Grim record (Levin et al., 2013). We also consider the Niwot Ridge, USA $\Delta^{14}CO_2$ record (NWR, 40.05°N, 105.59°W, 3523 m asl), which began in 2003 (Turnbull et al., 2007; Lehman et al., 2013). Niwot Ridge is also a mid-latitude high-altitude site, but samples are collected as whole air in flasks and measured by AMS in a similar manner to that described for the Wellington flask samples. Thus, we are comparing two independent Southern Hemisphere records with two independent Northern Hemisphere records, with the two hemispheres tied together by the common measurement laboratory used for Cape Grim and Jungfraujoch. Results from all records are compared in figure ~~7~~56.

The Wellington and Cape Grim records are generally consistent with one another (Figure 3), with the exception of the 1995-2005 period, when the Wellington record is slightly higher, apparently due to bias in the Wellington record (discussed in section 3.5.3.). Differences between the sites are smaller than the measurement uncertainty for all other periods (table 2). This implies that ~~the~~ $\Delta^{14}CO_2$ is ~~signal is homogenous~~homogeneous across Southern Hemisphere clean air sites within the same latitude band, at least since the 1980s when the two records overlap. Similarly, the high altitude, mid-latitude Northern Hemisphere sites are consistent with one another, although there are some differences in seasonal cycles in recent years (Turnbull et al., 2009b).

The bomb spike maximum is higher and earlier in the Northern Hemisphere records (figure 6), consistent with the production of most bomb ^{14}C in the Northern Hemisphere

stratosphere ~~(figure 5)~~. We make a new, simple estimate of the interhemispheric exchange time during the 1963 – 1965 period using the difference in the timing of the Northern and Southern Hemisphere bomb peaks. The first maximum ~~We determine a new estimate for the interhemispheric exchange time from the difference in timing of the first maximum~~ of the bomb peak ~~in each hemisphere (~~was in July 1963 in the Northern Hemisphere and January 1965 in the Southern Hemisphere, a ~~as~~ 1.4 ~~years~~year offset, implying a 1.4 year exchange time. This is consistent with other more detailed interhemispheric exchange time estimates that have been determined from long-term measurements of SF_6 of 1.3 to 1.4 years (Geller at el., 1997; Patra et al., 2011).

Northern Hemisphere $\Delta^{14}CO_2$ remains higher than Southern Hemisphere $\Delta^{14}CO_2$ by about 20 ‰ until 1972. Although most nuclear weapons testing ceased in 1963, a few smaller tests continued in the late 1960s, contributing to this continued interhemispheric offset (Enting, 1982). The interhemispheric gradient disappeared within about 1.5 years after atmospheric testing essentially stopped in 1970. Except periods of noisy data from Vermunt in the late 1970s and Wellington in 1995-2005, there are only small (<2 ‰) interhemispheric gradients from 1972 until 2002 (figure ~~7~~6, table 2).

As previously noted by Levin et al. (2010) using a shorter dataset, ~~From 2002,~~ an interhemispheric gradient of 5-7 ‰ develops in 2002, with the Southern Hemisphere sites higher than the Northern Hemisphere sites (table 2). We choose 1986 – 1990 and 2005 – 2013 as time periods to compare, to avoid the periods where the Wellington record is noisy (1995 – 2005) and where we substituted flask measurements from 1990 – 1993. In 1986 – 1990, there is less than 2 ‰ difference between Wellington and either Cape Grim or Jungfraujoch. There is also no difference between the Cape Grim and Jungfraujoch records during this time period. The Wellington and Cape Grim records still agree within 2 ‰ after 2005, but both Jungfraujoch and Niwot Ridge diverge from Wellington, by 4.8 ± 2.7 and 6.9 ± 2.5 ‰, respectively; ~~they~~ Jungfraujoch and Niwot Ridge are not significantly different from one another. This new interhemispheric gradient is robust, being consistent amongst the sites measured by three different research groups each with their own methods. It is not an artifact of interlaboratory offsets, since Cape Grim and Jungfraujoch measurements are made by the same group using the same sampling and measurement methods, and the Wellington and Niwot Ridge measurements (measured by different techniques) agree well with the other sites at similar latitude (Cape Grim and Jungfraujoch respectively). This developing gradient is also apparent in the larger sampling network of Levin et al (2010) ~~also apparent in~~and ~~2005 – 2007~~ in a separate $\Delta^{14}CO_2$ sampling network (Graven et al., 2012), although that dataset extends only to 2007.

Graven et al. (2012) demonstrated that increasing (mostly Northern Hemisphere) fossil fuel CO_2 emissions cannot explain this $\Delta^{14}CO_2$ interhemispheric gradient, and instead, they postulated that ^{14}C uptake into the Southern Ocean reduced over time. Levin et al. (2010) were able to roughly replicate this interhemispheric gradient in their GRACE model~~, which tunes~~ by tuning the terrestrial biosphere fluxes to match the observed global average atmospheric CO_2~~CO2~~ and $\Delta^{14}CO_2$ records. Where the observations suggest the rapid development of an interhemispheric gradient in the early 2000's, the

GRACE model simulates a more gradual transition ~~from~~ over a period of roughly two decades.

~~Given the limited spatial coverage of the current $\Delta^{14}CO_2$ observing network, it is not possible to robustly determine whether the interhemispheric gradient is due to terrestrial processes occurring predominantly in the Northern Hemisphere or ocean processes in the Southern Hemisphere from the existing $\Delta^{14}CO_2$ data alone.~~ Independent evidence suggests that the Southern Ocean is more likely to be responsible for this rapid shift in the atmospheric $\Delta^{14}CO_2$ gradient.

~~The development of the interhemispheric $\Delta^{14}CO_2$ gradient coincides with an apparent~~apparent That is, an apparent reorganization of Southern Ocean carbon exchange in the early 2000s (Landschützer et al., 2015)~~.~~ is postulated to be associated with changes in upwelling of deep water (DeVries et al., 2017), to which atmospheric $\Delta^{14}CO_2$ is highly sensitive (Rodgers et al., 2011; Graven et al., 2012b). The observed $\Delta^{14}CO_2$ interhemispheric gradient is consistent with these postulated changes in upwelling. ~~However~~Other possible explanations for this new interhemispheric $\Delta^{14}CO_2$ gradient are~~,~~ ~~but~~The net Southern Ocean carbon sink is determined by the balance between CO_2 uptake into surface waters, which are then subducted and sequester carbon, and release of carbon to the atmosphere from upwelling of very old, carbon-rich deep waters. CO_2 uptake into surface waters cannot change atmospheric $\Delta^{14}CO_2$, since the $\Delta^{14}C$ notation includes a mathematical correction for natural isotopic fractionation. In contrast, the ^{14}C disequilibrium between old (and therefore ^{14}C-poor), deep waters and the atmosphere means that release of CO_2 from the Southern Ocean to the atmosphere decreases atmospheric $\Delta^{14}CO_2$; the magnitude of that decrease depends on both the carbon flux and the ^{14}C disequilibrium. Thus, since the 1980s, atmospheric $\Delta^{14}C$ has been highly sensitive to Southern Ocean upwelling, the same mechanism that governs the ocean CO_2 sink (Graven et al., 2012). Model simulations suggest that changes in Southern Ocean ventilation may have played a key role in pre-industrial variations in the latitudinal gradient of atmospheric $^{14}CO_2$ (Rodgers et al., 2011).

~~Several studies using both data and modeling suggests that the climate-induced increase in westerly winds over the Southern Ocean increased upwelling of carbon-rich deep waters and thus reduced the Southern Ocean CO_2 sink efficiency (Le Quéré et al., 2007; Sitch et al., 2015). Yet, more recent evidence suggests a reinvigorated Southern Ocean carbon sink since about 2002 (Munro et al., 2016; Landschützer et al., 2015). These studies suggest that multiple factors contributed to the reinvigorated carbon sink, with different controls in the different Southern Ocean regions; these data support a decreasing upwelling of old, deep waters in recent years. Decreased upwelling would also cause a relative increase in Southern Hemisphere $\Delta^{14}CO_2$ and thus drive the observed interhemispheric $\Delta^{14}CO_2$ gradient, which appears at the same time as the apparent reinvigoration of the carbon sink in the early 2000s.~~

~~Although the changing Southern Ocean carbon sink is the most likely explanation,~~substantial underreporting of Northern Hemisphere fossil CO_2 emissions (e.g. Francey et al., 2013) or changes in the land carbon sink (Wang et al., 2013; Sitch et al.,

2015~~; Wang et al., 2013~~) ~~could also explain the new interhemispheric $\Delta^{14}CO_2$ gradient~~.
~~More observations of the spatial and temporal variations of atmospheric $\Delta^{14}CO_2$ are critical to unlocking this trace gas's full potential to inform us about the processes controlling the global carbon cycle.~~ Given the limited spatial coverage of the current $\Delta^{14}CO_2$ observing network, it is not possible to robustly determine ~~whether~~which of these processes causes the interhemispheric gradient ~~is due to terrestrial processes occurring predominantly in the Northern Hemisphere or ocean processes in the Southern Hemisphere from the existing $\Delta^{14}CO_2$ data alone~~. This could be achieved with more observations of the spatial and temporal variations of atmospheric $\Delta^{14}CO_2$.

Comment [SM2]: Maybe a bit too on the nose? Delete if you think so.

5. Conclusions

The 60 year-long Wellington $\Delta^{14}CO_2$ record has been revised and extended to 2014. Most revisions were minor, but we particularly note that the earlier reported 1990-1993 measurements have been entirely replaced with new measurements. A second period form 1995-2005 has poorer data quality than the rest of the record, and may also be biased high by a few permil. These data have been revised substantially, and new measurements have been added to this period, but we were unable to definitively identify or correct for bias, so the data have been retained, albeit with caution. We further validated the record by comparison with tree ring samples collected from the Baring Head sampling location and from nearby Eastbourne, Wellington; both tree ring records show excellent agreement with the original record, and indicate that there are no other periods where the original measurements are problematic.

The Wellington $\Delta^{14}CO_2$ time series records the history of atmospheric nuclear weapons testing and the subsequent decline of $\Delta^{14}CO_2$ as the bomb ^{14}C moved throughout the carbon cycle, and ^{14}C-free fossil fuel emissions further decreased $\Delta^{14}CO_2$. The timing of the first appearance of the bomb-^{14}C peak at Wellington is consistent with other recent estimates of interhemispheric exchange time at 1.4 years.

The seasonal cycle at Wellington evolves through the record, apparently dominated by the seasonality of cross-tropopause transport, which drives a changing seasonal cycle through time. In the early post-bomb period, the seasonally variable movement of bomb ^{14}C from the Northern Stratosphere through the Northern Troposphere to the Southern Troposphere appears to be the dominant control on the seasonal cycle at Wellington. ~~Southern Hemisphere troposphere was enriched in ^{14}C relative to the Southern Hemisphere stratosphere so that the seasonal minimum occurred at Wellington when cross-tropopause transport is at a maximum.~~ The seasonal cycle reversed ~~once the bomb perturbation reduced and continuing natural cosmogenic production meant that the Southern Hemisphere stratosphere was once again enriched in ^{14}C relative to the troposphere~~in later years, possibly due to a change in sign of the terrestrial biosphere $\Delta^{14}C$ signal. In recent years, the seasonal cycle has an amplitude of only 2 ‰, with a maximum in the austral spring. Cape Grim exhibits a similar seasonal cycle magnitude, but appears to be very slightly influenced by a terrestrial/anthropogenic signal during the austral winter that is not apparent at Wellington.

During the 1980s and 1990s, $\Delta^{14}CO_2$ was similar at mid-latitude clean air sites in both hemispheres, but since the early 2000s, the Northern Hemisphere $\Delta^{14}CO_2$ has dropped below the Southern Hemisphere by 5-7 ‰. The control on this changing interhemispheric gradient cannot be robustly determined from the existing sparse $\Delta^{14}CO_2$ observations, ~~This is most likely due to a change~~but may be due to a change in Southern Ocean dynamics reducing upwelling of old, ^{14}C-poor deep waters, ~~which is~~ consistent with recent evidence for an increasing Southern Ocean carbon sink. This ~~result~~ implies that ongoing and expanded Southern Hemisphere $\Delta^{14}CO_2$ observations and modelling ~~can~~ may provide a fundamental constraint on our understanding of Southern Ocean dynamics and exchange processes.

6. Acknowledgements

A 60 year-long record takes more than a handful of authors to produce. This work was possible only because of the amazing foresight and scientific understanding of Athol Rafter and Gordon Fergusson, who began this record in the 1950s. Their work was continued over the years by a number of people, including Hugh Melhuish, Martin Manning, Dave Lowe, Rodger Sparks, Charlie McGill, Max Burr and Graeme Lyon. This work was funded by the Government of New Zealand as GNS Science Global Change Through Time core funding and NIWA Greenhouse Gases, Emissions, and Carbon Cycle Science Programme core funding. The author(s) wish to acknowledge the contribution of New Zealand eScience Infrastructure (NeSI) to the results of this research. New Zealand's national computer~~compute~~ and analytics services and team are supported by the NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation and Employment (http://www.nesi.org.nz). We thank Dr Scott Lehman (University of Colorado) and Dr Ingeborg Levin (University of Heidelberg) for sharing their $\Delta^{14}CO_2$ datasets for comparison with the Wellington record.

7. Data availability

The datasets presented in this paper are included as supplementary material. The datasets (including updates as they are available) can be accessed through the World Data Centre for Greenhouse Gases (http://ds.data.jma.go.jp/gmd/wdcgg/) or directly through GNS Science (https://gns.cri.nz/Home/Products/Databases/Wellington-atmospheric-14CO2-record) or NIWA (ftp://ftp.niwa.co.nz/tropac/).

7. References

AECOM New Zealand Limited, 2016. Community greenhouse gas inventory for Wellington City and the Greater Wellington Region 2000-2015, Wellington.

Australian Government, 2016. State and territory greenhouse gas inventories 2014. Department of the Environment.

Baisden, W.T., Prior, C.A., Chambers, D., Canessa, S., Phillips, A., Bertrand, C., Zondervan, A., Turnbull, J.C., 2013. Radiocarbon sample preparation and data flow at Rafter: Accommodating enhanced throughput and precision. Nuclear Instruments and Methods B294, 194-198.

Bevington, P.R., Robinson, D.K., 2003. Data reduction and error analysis for the physical sciences, Third Edition. McGraw-Hill.

Boden, T.A., Marland, G., Andres, R.J., 2012. Global, Regional, and National Fossil-Fuel CO_2 Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Bozhinova, D., Combe, M., Palstra, S.W.L., Meijer, H.A.J., Krol, M.C., Peters, W., 2013. The importance of crop growth modeling to interpret the $^{14}CO_2$ signature of annual plants. Global Biogeochemical Cycles 27, 792-803.

Brailsford, G.W., Stephens, B.B., Gomez, A.J., Riedel, K., Mikaloff Fletcher, S.E., Nichol, S.E., Manning, M.R., 2012. Long-term continuous atmospheric CO_2 measurements at Baring Head, New Zealand. Atmospheric Measurement Techniques 5, 3109-3117.

Broecker, W.S., Peng, T.-H., Ostlund, H., Stuiver, M., 1985. The distribution of bomb radiocarbon in the ocean. Journal of Geophysical Research C4, 6953-6970.

Caldeira, K., Rau, G.H., Duffy, P.B., 1998. Predicted net efflux of radiocarbon from the ocean and increase in atmospheric radiocarbon content. Geophysical Research Letters 25, 3811-3814.

Cleveland, R., Cleveland, W., McRae, J., Terpenning, I., 1990. STL: A seasonal-trend decomposition procedure based on Loess. Journal of Official Statistics 6, 3-33.

Currie, K.I., Brailsford, G., Nichol, S., Gomez, A., Sparks, R., Lassey, K.R., Riedel, K., 2011. Tropospheric $^{14}CO_2$ at Wellington, New Zealand: the world's longest record. Biogeochemistry 104, 5-22.

Davies, T., Cullen, M., Malcolm, A., Mawson, M., Staniforth, A., White, A., Wood, N., 2005. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. Quarterly Journal of the Royal Meteorological Society 131, 1759-2005.

DeVries, T., Holzer, M., Primeau, F., 2017. Recent increase in oceanic carbon uptake driven by weaker upper-ocean overturning. Nature 542, 215-218.

Djuricin, S., Pataki, D.E., Xu, X., 2010. A comparison of tracer methods for quantifying CO_2 sources in an urban region. Journal of Geophysical Research 115.

Enting, I.G., 1982. Nuclear weapons data for use in carbon cycle modelling. CSIRO Division of Atmospheric Physics and Technology, Melbourne, Australia.

Ferretti, D.F., Lowe, D.C., Martin, R.H., Brailsford, G.W., 2000. A new gas chromatograph-isotope ratio mass spectrometry technique for high-precision, N_2O-free analysis of $\delta^{13}C$ and $\delta^{18}O$ in atmospheric CO_2 from small air samples. Journal of Geophysical Research Atmospheres 105, 6709-6718.

855 Francey, R.J., Trudinger, C.M., van der Schoot, M., Law, R.M., Krummel, P.B.,
856 Langenfelds, R.L., Steele, L.P., Allison, C.E., Stavert, A.R., Andres, R.J.,
857 Rödenbeck, C., 2013. Atmospheric verification of anthropogenic CO_2 emission
858 trends. Nature Climate Change 3, 520-524.
859 Geller, L.S., Elkins, J.W., Lobert, J.M., Clarke, A.D., Hurst, D.F., Butler, J.H., Myers,
860 R.C., 1997. Tropospheric SF_6: Observed latitudinal distribution and trends, derived
861 emissions and interhemispheric exchange time. Geophysical Research Letters 24,
862 675-678.
863 Graven, H.D., Gruber, N., 2011. Continental-scale enrichment of atmospheric $^{14}CO_2$ from
864 the nuclear power industry: potential impact on the estimation of fossil fuel-derived
865 CO_2. Atmospheric Chemistry and Physics 11, 12339-12349.
866 Graven, H.D., Guilderson, T.P., Keeling, R.F., 2012. Observations of radiocarbon in CO_2
867 at seven global sampling sites in the Scripps flask network: Analysis of spatial
868 gradients and seasonal cycles. Journal of Geophysical Research 117.
869 Hogg, A.G., 2013. SHCAL13 Southern Hemisphere calibration, 0-50,000 years CAL BP.
870 Radiocarbon.
871 Hua, Q., Barbetti, M., Jacobsen, G., Zoppi, U., Lawson, E., 2000. Bomb radiocarbon in
872 annual tree rings from Thailand and Australia. Nuc. Inst. and Meth. in Physics
873 Research B 172, 359-365.
874 Hua, Q., Barbetti, M., Rakowski, A.Z., 2013. Atmospheric radiocarbon for the period
875 1950-2010. Radiocarbon 55, 1-14.
876 Jones, A.R., Thomson, D., Hort, M., Devenish, B., 2007. The UK Met Office's next-
877 generation atmospheric dispersion model, NAME III, in: Borrego, C., Norman, A.-
878 L. (Eds.), Air Pollution Modeling and Its Application XVII. Springer.
879 Kanu, A., Comfort, L., Guilderson, T.P., Cameron-Smith, P.J., Bergmann, D.J., Atlas,
880 E.L., Schauffler, S., Boering, K.A., 2015. Measurements and modelling of
881 contemporary radiocarbon in the stratosphere. Geophysical Research Letters 43.
882 Keeling, C.D., Piper, S.C., Whorf, T.P., Keeling, R.F., 2011. Evolution of natural and
883 anthropogenic fluxes of atmospheric CO2 from 1957 to 2003. Tellus B 63, 1-22.
884 Keeling, C.D., Whorf, T., 2005. Atmospheric CO_2 records from sites in the SIO air
885 sampling network, Trends: A compendium of data of global change. Carbon
886 Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge,
887 Tenn., USA.
888 Key, R.M., 2004. A global ocean carbon climatology: Results from Global Data Analysis
889 Project (GLODAP). Global Biogeochemical Cycles 18.
890 Kjellström, E., Feichter, J., Hoffman, G., 2000. Transport of SF_6 and $^{14}CO_2$ in the
891 atmospheric general circulation model ECHAM4. Tellus 52B, 1-18.
892 Landschützer, P., Gruber, N., Haumann, F.A., Rödenbeck, C., Bakker, D.C.E., van
893 Heuven, S., Hoppema, M., Metzl, N., Sweeney, C., Takahashi, T., Tilbrook, B.,
894 Wanninkhof, R., 2015. The reinvigoration of the Southern Ocean carbon sink.
895 Science 349, 1221-1224.
896 Law, R.M., Steele, L.P., Krummel, P.B., Zahorowski, W., 2010. Synoptic variations in
897 atmospheric CO_2 at Cape Grim: a model intercomparison. Tellus B 62, 810-820.
898 Le Quere, C., Rodenbeck, C., Buitenhuis, E.T., Conway, T.J., Langenfelds, R., Gomez,
899 A., Labuschagne, C., Ramonet, M., Nakazawa, T., Metzl, N., Gillett, N., Heimann,

M., 2007. Saturation of the Southern Ocean CO_2 Sink Due to Recent Climate Change. Science 316, 1735-1738.

Lehman, S.J., Miller, J.B., Wolak, C., Southon, J.R., Tans, P.P., Montzka, S.A., Sweeney, C., Andrews, A.E., LaFranchi, B.W., Guilderson, T.P., Turnbull, J.C., 2013. Allocation of terrestrial carbon sources using $^{14}CO_2$: Methods, measurement, and modelling. Radiocarbon 55, 1484-1495.

Levin, I., Kromer, B., Hammer, S., 2013. Atmospheric $\Delta^{14}CO_2$ trend in Western European background air from 2000 to 2012. Tellus B 65.

Levin, I., Kromer, B., Schmidt, M., Sartorius, H., 2003. A novel approach for independent budgeting of fossil fuel CO_2 over Europe by $^{14}CO_2$ observations. Geophysical Research Letters 30, 2194.

Levin, I., Kromer, B., Schoch-Fischer, H., Bruns, M., Munnich, M., Berdau, D., Vogel, J.C., Munnich, K.O., 1985. 25 years of tropospheric ^{14}C observations in central Europe. Radiocarbon 27, 1-19.

Levin, I., Naegler, T., Kromer, B., Diehl, M., Francey, R.J., Gomez-Pelaez, A.J., Steele, L.P., Wagenbach, D., Weller, R., Worthy, D.E., 2010. Observations and modelling of the global distribution and long-term trend of atmospheric $^{14}CO_2$. Tellus B 62, 26-46.

Lopez, M., Schmidt, M., Delmotte, M., Colomb, A., Gros, V., Janssen, C., Lehman, S.J., Mondelain, D., Perrussel, O., Ramonet, M., Xueref-Remy, I., Bousquet, P., 2013. CO, NOx and $^{13}CO_2$ as tracers for fossil fuel CO_2: results from a pilot study in Paris during winter 2010. Atmospheric Chemistry and Physics 13, 7343-7358.

Lowe, D.C., Judd, W., 1987. Graphite target preparation for radiocarbon dating by accelerator mass spectrometry. Nuclear Instruments and Methods in Physics Research B 28, 113-116.

Manning, M.R., Lowe, D.C., Melhuish, W.H., Sparks, R.J., Wallace, G., Brenninkmeijer, C.A.M., McGill, R.C., 1990. The use of radiocarbon measurements in atmospheric sciences. Radiocarbon 32, 37-58.

Meijer, H.A.J., Pertuisot, M.-H., van der Plicht, J., 2006. High accuracy ^{14}C measurements for atmospheric CO_2 samples by AMS. Radiocarbon 48, 355-372.

Meijer, H.A.J., Smid, H.M., Perez, E., Keizer, M.G., 1996. Isotopic characterization of anthropogenic CO_2 emissions using isotopic and radiocarbon analysis. Physical Chemistry of the Earth 21, 483-487.

Miller, J.B., Lehman, S.J., Montzka, S.A., Sweeney, C., Miller, B.R., Wolak, C., Dlugokencky, E.J., Southon, J.R., Turnbull, J.C., Tans, P.P., 2012. Linking emissions of fossil fuel CO_2 and other anthropogenic trace gases using atmospheric $^{14}CO_2$. Journal of Geophysical Research 117, D08302.

Munro, D.R., Lovenduski, N.S., Takahashi, T., Stephens, B.B., Newberger, T., Sweeney, C., 2015. Recent evidence for a strengthening CO2sink in the Southern Ocean from carbonate system measurements in the Drake Passage (2002-2015). Geophysical Research Letters, n/a-n/a.

Naegler, T., Ciais, P., Rodgers, K., Levin, I., 2006. Excess radiocarbon constraints on air-sea gas exchange and the uptake of CO_2 by the oceans. Geophysical Research Letters 33.

Naegler, T., Levin, I., 2009. Observation-based global biospheric excess radiocarbon inventory 1963–2005. Journal of Geophysical Research 114.

946 Norris, M.W., 2015. Reconstruction of historic fossil CO_2 emissions using radiocarbon
947 measurements from tree rings, School of Geography, Environment and Earth
948 Sciences. Victoria University of Wellington.
949 Nydal, R., Lövseth, K., 1983. Tracing bomb ^{14}C in the atmosphere 1962-1980. Journal of
950 Geophysical Research 88, 3621-3642.
951 Oeschger, H., Siegenthaler, U., Schotterer, U., Gugelmann, A., 1975. A box diffusion
952 model to study the carbon dioxide exchange in nature. Tellus XXVII, 168-192.
953 Otago Daily Times, 1957. Polar ice caps may melt with industrialisation, Otago Daily
954 Times, 23/1/1957 ed, Dunedin, New Zealand, p. 1.
955 Patra, P.K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H.,
956 Cameron-Smith, P., Chipperfield, M.P., Corbin, K., Fortems-Cheiney, A., Fraser,
957 A., Gloor, E., Hess, P., Ito, A., Kawa, S.R., Law, R.M., Loh, Z., Maksyutov, S.,
958 Meng, L., Palmer, P.I., Prinn, R.G., Rigby, M., Saito, R., Wilson, C., 2011.
959 TransCom model simulations of CH_4 and related species: linking transport, surface
960 flux and chemical loss with CH_4 variability in the troposphere and lower
961 stratosphere. Atmospheric Chemistry and Physics 11, 12813-12837.
962 Pickers, P.A., Manning, A.C., 2015. Investigating bias in the application of curve fitting
963 programs to atmospheric time series. Atmospheric Measurement Techniques 8,
964 1469-1489.
965 Rafter, T.A., 1955. ^{14}C variations in nature and the effect on radiocarbon dating. New
966 Zealand Journal of Science and Technology B37.
967 Rafter, T.A., Fergusson, G., 1959. Atmospheric radiocarbon as a tracer in geophysical
968 circulation problems, United Nations Peaceful Uses of Atomic Energy. Pergamon
969 Press, London.
970 Randerson, J.T., Enting, I.G., Schuur, E.A.G., Caldeira, K., Fung, I.Y., 2002. Seasonal
971 and latitudinal variability of troposphere $\Delta^{14}CO_2$: Post bomb contributions from
972 fossil fuels, oceans, the stratosphere, and the terrestrial biosphere. Global
973 Biogeochemical Cycles 16, 1112.
974 Reimer, P.J., Brown, T.A., Reimer, R.W., 2004. Discussion: Reporting and calibration of
975 post-bomb ^{14}C data. Radiocarbon 46, 1299-1304.
976 Rodgers, K.B., Mikaloff-Fletcher, S.E., Bianchi, D., Beaulieu, C., Galbraith, E.D.,
977 Gnanadesikan, A., Hogg, A.G., Iudicone, D., Lintner, B.R., Naegler, T., Reimer,
978 P.J., Sarmiento, J.L., Slater, R.D., 2011. Interhemispheric gradient of atmospheric
979 radiocarbon reveals natural variability of Southern Ocean winds. Climate of the
980 Past 7, 1123-1138.
981 Sitch, S., Friedlingstein, P., Gruber, N., Jones, S.D., Murray-Tortarolo, G., Ahlström, A.,
982 Doney, S.C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P.E., Lomas,
983 M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L.,
984 Canadell, J.G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S.L.,
985 Le Quéré, C., Smith, B., Zhu, Z., Myneni, R., 2015. Recent trends and drivers of
986 regional sources and sinks of carbon dioxide. Biogeosciences 12, 653-679.
987 Steinkamp, K., Mikaloff Fletcher, S.E., Brailsford, G., Smale, D., Moore, S., Keller, E.D.,
988 Baisden, W.T., Mukai, H., Stephens, B.B., 2016. Atmospheric CO_2 observations
989 and models suggest strong carbon uptake by forests in New Zealand. Atmospheric
990 Chemistry and Physics Discussions, 1-55.

Stephens, B.B., Brailsford, G.W., Gomez, A.J., Riedel, K., Mikaloff Fletcher, S.E., Nichol, S., Manning, M., 2013. Analysis of a 39-year continuous atmospheric CO_2 record from Baring Head, New Zealand. Biogeosciences 10, 2683-2697.

Stuiver, M., Polach, H.A., 1977. Discussion: Reporting of ^{14}C data. Radiocarbon 19, 355-363.

Stuiver, M., Quay, P.D., 1981. Atmospheric 14C changes resulting from fossil fuel CO2 release and cosmic ray flux variability. Earth and Planetary Science Letters, 53, 349-362.

Suess, H.E., 1955. Radiocarbon concentration in modern wood. Science 122, 414-417.

Sweeney, C., Gloor, E., Jacobson, A.R., Key, R.M., McKinley, G., Sarmiento, J.L., Wanninkhof, R., 2007. Constraining global air-sea gas exchange for CO_2 with recent bomb ^{14}C measurements. Global Biogeochemical Cycles 21.

Tans, P.P., De Jong, A.F., Mook, W.G., 1979. Natural atmospheric ^{14}C variation and the Suess effect. Nature 280, 826-828.

Thoning, K.W., Tans, P.P., Komhyr, W.D., 1989. Atmospheric carbon dioxide at Mauna Loa Observatory 2. Analysis of the NOAA GMCC data, 1974-1985. Journal of Geophysical Research 94, 8549-8563.

Trumbore, S.E., 2000. Age of soil organic matter and soil respiration: Radiocarbon constraints on belowground C dynamics. Ecological Applications 10, 399-411.

Turnbull, J.C., 2006. Development of a high precision $^{14}CO_2$ measurement capability and application to carbon cycle studies, Geological Sciences. University of Colorado, Boulder, p. 132.

Turnbull, J.C., Miller, J.B., Lehman, S.J., Hurst, D.F., Peters, W., Tans, P.P., Southon, J.R., Montzka, S.A., Elkins, J.W., Mondeel, D.J., Romashkin, P.A., Elansky, N.F., Shkorokhod, A., 2009. Spatial distribution of $\Delta^{14}CO_2$ across Eurasia: Measurements from the TROICA-8 expedition. Atmospheric Chemistry and Physics 9, 175-187.

Turnbull, J.C., Rayner, P.J., Miller, J.B., Naegler, T., Ciais, P., Cozic, A., 2009. On the use of $^{14}CO_2$ as a tracer for fossil fuel CO_2: quantifying uncertainties using an atmospheric transport model. Journal of Geophysical Research 114, D22302.

Turnbull, J.C., Sweeney, C., Karion, A., Newberger, T., Lehman, S.J., Tans, P.P., Davis, K.J., Lauvaux, T., Miles, N.L., Richardson, S.J., Cambaliza, M.O., Shepson, P.B., Gurney, K., Patarasuk, R., Razlivanov, I., 2015. Toward quantification and source sector identification of fossil fuel CO2 emissions from an urban area: Results from the INFLUX experiment. Journal of Geophysical Research: Atmospheres.

Turnbull, J.C., Zondervan, A., Kaiser, J., Norris, M., Dahl, J., Baisden, W.T., Lehman, S.J., 2015. High-precision atmospheric $^{14}CO_2$ measurement at the Rafter Radiocarbon Laboratory. Radiocarbon 57, 377-388.

Wang, Y., Li, M., Shen, L., 2013. Accelerating carbon uptake in the Northern Hemisphere: evidence from the interhemispheric difference of atmospheric CO_2 concentrations. Tellus B 65.

Ziehn, T., 2014. Greenhouse gas network design using backward Lagrangian particle dispersion modelling – Part 1: Methodology and Australian test case. Atmospheric Chemistry and Physics 14.

Zondervan, A., Hauser, T., Kaiser, J., Kitchen, R., Turnbull, J.C., West, J.G., 2015. XCAMS: The compact ^{14}C accelerator mass spectrometer extended for 10Be and

26Al at GNS Science, New Zealand. Nuclear Instruments and Methods B361, 25-33.
Zondervan, A., Sparks, R.J., 1996. Development plans for the AMS facility at the Institute of Geological and Nuclear Sciences, New Zealand. Radiocarbon 38, 133-134.